DOCUMENT RESUME

ED 077 270 FL 003 849

AUTHOR	Cherrier, Roy, Ed.
TITLE	English Language Testing. Report of the RELC Fifth Regional Seminar, Bangkok, 25-30 May, 1970.
INSTITUTION	Regional English Center (Singapore).
PUB DATE	71
NOTE	232p.
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	Admission Criteria; Conference Reports; Educational Improvement; *English (Second Language); Error Patterns; *Language Proficiency; *Language Skills; Language Tests; Predictive Ability (Testing); Second Language Learning; *Seminars; Test Construction; *Testing; Test Reliability

ABSTRACT

The primary objective of the Regional English Language Center (RELC) is the improvement of standards of teaching and learning English as a second or foreign language in the countries of Southeast Asia: Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand, and Vietnam. This report on the RELC seminar on English language testing contains the papers presented at the plenary sessions of the meeting as well as summaries of the discussion from each session; also included are supplementary papers and reports from the workshops held during the meeting. Topics include a variety of issues in English language testing from general subjects, such as the present status of language testing in the member countries and the role of contrastive analysis in English language testing, to more specific remarks on English language testing, university entrance requirements, and papers on the testing of particular language skills. (VM)

ED 077270

ENGLISH LANGUAGE TESTING

Report of the RELC Fifth Regional Seminar

Bangkok, 25—30 May, 1970

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FL 003 849

Roy Cherrier, Editor

Southeast Asian Ministers of Education Organization
Regional English Language Centre
Singapore, 1971

CONTENTS

Page

Foreword - - - - - - - - - - - - - - i

Opening Ceremony

Welcoming Address: Tai Yu-lin - - - - - - - - - 1

Address: Gen. Netr Khemayodhin - - - - - - - - - 4

Opening Address: H.E. Dr. Sukich Nimmanheminda - - - 6

Messages by Non-SEAMEO government representatives:

United Kingdom - - - - - - - - - - 7
United States - - - - - - - - - - 7
Canada - - - - - - - - - - - 8
Hong Kong - - - - - - - - - - 9
India - - - - - - - - - - - - 10

Plenary Session I:

"Elements to Test in English as a Second Language", Bonifacio Sibayan - - - - - - - - - - - - - 11

Discussion - - - - - - - - - - - - 22

Plenary Session II:

"Panel Discussion on Objectives of English Teaching and Present Status of English Language Testing in SEAMEO Countries" - 23

Indonesia: P. Tirtopramono - - - - - - - - 23
Laos: Pinhkham Simmalavong - - - - - - - 28
Malaysia: Murad bin Mohd. Nor - - - - - - - 34
Philippines: Clodoaldo H. Leocadio - - - - - - 40
Singapore: Molly Ng Kuen Seong - - - - - - 44
Thailand: Banlue Tinpanga - - - - - - - - 51
Vietnam: Duong Thanh Binh - - - - - - - 55

Discussion - - - - - - - - - - - - - 67

Plenary Session III:

"Context for Language Testing" John A. Upshur - - - - - 73

Discussion - - - - - - - - - - - - - 84

CONTENTS—*contd.*

Page

Plenary Session IV:

"Contrastive Analysis and English Language Testing" Gerhard Nickel - - - - - - - - - - - - - 85

Discussion - - - - - - - - - - - - - 94

Plenary Session V:

"Language Tests and Language Errors" Alan Davies - - - 95

Discussion - - - - - - - - - - - - - 104

Plenary Session VI:

"Error Analysis as an Index of Class Progress " Lim Kiat Boey - 105

Discussion - - - - - - - - - - - - - 109

Plenary Session VII:

"English Language Testing for University Entrance", Robert B. Kaplan - - - - - - - - - - - - - 113

Discussion - - - - - - - - - - - - - 118

"Standardized English Tests for University Entrance", M. Moehnilabib - - - - - - - - - - - - - 119

Plenary Session VIII:

"Predicting Individual Differences in English Language Learning" John B. Carroll - - - - - - - - - - - - - 125

Discussion - - - - - - - - - - - - - 132

Plenary Session IX:

"Discrete versus Non-Discrete Testing in English Language Proficiency", Eugene J. Briere - - - - - - - - - - 133

Discussion - - - - - - - - - - - - - 139

Plenary Session X:

"Translation in Language Testing", Huynh Dinh Te - - - 141

Discussion - - - - - - - - - - - - - 144

Supplementary Papers:

"English Language Testing Targets and Techniques", S. N. Saha - 145

"Some Problems of Validity in Language Testing", P. W. J. Nababan - - - - - - - - - - - - - 157

"Testing Reading Skills", B. Deane Smith - - - - - - - 166

"A Research Philosophy and English Language Testing", Peter B. Botsman - - - - - - - - - - - - - 175

"Objectives, Learning Tasks, and Testing in Foreign Languages"', Eric W. Bauer - - - - - - - - - - - - - 182

CONTENTS—*contd.*

Page

Workshop Reports:

1. "Testing Listening Skills", Ralph P. Barrett - - - - - 191
2. "Testing Oral Skills", Charles R. Sheehan - - - - - 193
3. "Testing Reading Skills", B. Deane Smith - - - - - 195
4. "Testing Writing Skills", Oliver Seet Beng Hean - - - 197
5. "Testing Selected Areas of English Grammar", Eric W. Bauer - 199
6. "Setting Realistic Behavioral Objectives for Determining Progress in Language Mastery at Various Levels", Richard Noss - - - - - - - - - - - - 203
7. "Testing Foreign Language Aptitude", John B. Carroll - 206
8. "Techniques of Test Construction", Alan Davies - - - 209
9. "Administrative Problems of English Language Testing", Murad bin Mohd. Nor - - - - - - - - - - 210
10. "Preparing an English Proficiency Test for English Teacher Applicants in South East Asia", Joe. D. Palmer - - - 213

Closing Ceremony

Address on behalf of the Ministry of Education, Thailand - - 215
Closing Address: Tai Yu-lin - - - - - - - - - 216

Seminar Officials - - - - - - - - - - - - - 219
Speakers and Workshop Leaders - - - - - - - - 220
Participants - - - - - - - - - - - - - 222

FOREWORD

The Regional English Language Centre (RELC), located in Singapore, is one of the educational centres developed under the South East Asian Ministers of Education Organization (SEAMEO). Its primary objective is the improvement of standards of teaching and learning English as a second or foreign language in the countries of Southeast Asia. Its activities include training, research, preparation of instructional materials, consultant and exchange services, and information services. One of its specific activities is the sponsorship of seminars on pertinent topics in English language teaching.

The Seminar on English Language Testing, here reported, is the fifth in the RELC seminar series. The first, held in Singapore in 1967, formulated development plans for the Centre. The second, held in April 1968, dealt with the training of English language teachers in the SEAMEO region, and the third, held in October of the same year, analyzed and evaluated the results of the Teacher Education Programme tryouts. These first three were limited in participation, and dealt with specific topics relevant to the formulation of policy and procedures for RELC activity. The fourth regional seminar, however, inaugurated the series of professional seminars of international scope. Held in Singapore in June 1969, on the topic "New Developments in the Theory and Methods of Teaching and Learning English", it drew 200 participants from 16 countries. Many of the papers presented at this seminar have been published in the *RELC Journal* (June 1970).

The RELC fifth Regional Seminar might make several claims to distinction. With over 400 participants from 18 countries, it is quite possibly the largest seminar ever devoted to the topic of English language testing. More significantly, it benefitted from a sharp focus on urgent problems of a specific area, Southeast Asia. The presentation of the current language testing situation in the seven SEAMEO countries during the second plenary session of the seminar established the framework for discussions and workshop sessions. As can be seen, however, this regional focus implied no lack of variety in language testing problems discussed; the region has some problems in common, but each country also has its own individual English language teaching and testing situation.

This *Report* reproduces all papers and addresses delivered in plenary sessions of the seminar, the reports of each workshop, and a limited selection of other papers presented. Exigencies of space have obliged us to omit most of the materials presented or developed in the workshops, and also to report the plenary session discussions in severely summary form. We regret these enforced limitations.

Even more we regret that this report in no way conveys what was for most of us most memorable; that is, the spirit of the seminar, which enveloped intense cooperative effect in an atmosphere of high friendliness, mutual respect, and generous Thai hospitality.

Address by TAI YU-LIN

Director, SEAMEO Regional English Language Centre

Your Excellency, Mr. Director, Distinguished Delegates, Ladies and Gentlemen:

On behalf of the Regional English Language Centre in Singapore, I have great pleasure in welcoming this large assemblage of scholars to our regional seminar on English Language Testing. First and foremost I should like to express my sincere gratitude to His Excellency Dr. Sukich Nimmanheminda, Minister of Education of Thailand, for coming here to honor us by declaring open this important meeting in South East Asia. I am deeply grateful to His Excellency for giving us his valuable time as I know how busy he must be in his state responsibilities after the meeting of the South East Asian Ministers of Education Council held in Bangkok barely ten days ago. I should like to thank the Royal Thai government through the Ministry of Education for being the generous host to this important seminar. I am sure I am voicing the unanimous feeling of the participants in conveying our appreciation to the Royal Thai Government for the magnificent organization and the warm hospitality that is already so fully expressed.

Our work in preparing for this seminar has only been made possible by the good will and hard work of the organizing committee of Thailand, headed by Khun Runjuan Intarakamhang, with our very good friend, M.L. Boonlua Debyasuvarn as advisor and Mrs. Mayuri Sukwiwat as secretary. The Regional English Language Centre in Singapore is indebted to the Bangkok organizing committee for their wholehearted cooperation and vigorous assistance in the organization of this seminar of international scope. We also appreciate very much the courtesies and kindnesses extended to the seminar participants by the staff members of the University of Pittsburgh Project in Thailand, by the British Embassy and the British Council, and by the American University Alumni Language Center. I would also like to express my warm welcome and thanks to the many important guests who have endorsed the importance of this seminar by being present at the opening ceremonies.

My grateful thanks are due to Gen. Netr Khemayodhin, Director of the South East Asian Ministers of Education Secretariat, for his warm message in the souvenir program and his all-out support to make this major RELC event this year a successful one. It is also my great privilege and pleasure to extend a warm welcome to the distinguished participants who have come to this seminar from more than seventeen countries. In addition to strong teams of senior educational administrators and teaching and testing specialists from the seven SEAMEO member countries — namely Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand and Vietnam — we have distinguished participants from Australia, Canada, Germany, Hong Kong, India, Japan, Nepal, New Guinea, the United Kingdom and the United States, representing governments, university departments, professional organizations, research centres, foundations and international bodies. Many have travelled long distances to be with us and we are particularly grateful to them for accepting our invitation. Included are leading authorities on language testing whose names are associated with well-known texts on the subject or language tests that have won world-wide recognition. Many are from professional testing services where important and fruitful research is being undertaken in language testing. I should like to take this opportunity to once more thank Their Excellencies the Ministers of Education of our member countries for sending as official particiants the distinguished members of the RELC Coordinating Committee which is in fact the Centre's governing body. The Coordinating Committee Members are policy makers of the RELC, and represent their respective Ministers of Education at our committee meetings which are held at least once a year. They have also been co-planners of this seminar.

It is of particular gratification to me to be able to welcome some 280 participants from Thailand — from the Ministry of Education, university departments, teachers colleges and other institutions — who have been especially released to attend this seminar at a very demanding time of the academic year in Thailand.

This is the fifth regional seminar of the Regional English Language Centre. We are in the second year of our operational activities. Our fourth seminar, on "New Developments in the Theory and Methods of Teaching and Learning English", held in Singapore in June last year, was attended by over 200 participants from sixteen countries. We are very much encouraged to see so many of our friends who were with us last year back with us this year. We are certain that their support will once more generate important contributions which will go a long way to solving some of the pressing educational problems concerning the teaching of English as a second or foreign language in the region.

We in South East Asia are acutely aware that effective foreign language classroom techniques at various levels can only be devised and implemented if they are accompanied by adequate test instruments for selection and placement, for identifying problem areas of the learners and for determining the individual learner's progress and achievement. In the report on the conference on problems in foreign language testing held at the University of Michigan in 1967 it was stated that the idea for that meeting "grew out of private and public expression of dissatisfaction by a number of people who are involved in test production or who depend on test results, dissatisfaction with the quality and range of foreign language tests available, with the uses to which tests often are put, and with the ends they are made to serve. In translating the idea to plan, it became necessary to find representatives of language test writers, language test users, and language test needers who could all be available for a meeting at the same time".

Similar problems are faced by our member countries in South East Asia. Our problems are even more urgent, in view of the examination consciousness at all levels in many of our educational systems. Whether we like it or not, success in examinations is a dominant goal in the teaching and learning effort. If examination instruments and standards are unrealistic and inadequate, if language examinations do not accurately evaluate the desired objectives of the language program, then inevitably they will constrain and warp the learning effort to their own measure. We may preach all we like about the four language skills, or about a well-rounded grasp of all aspects of the language. The learner, with the innocent pragmatism of youth, translates it all into one simple question: "Will I be examined on that?" It is therefore fitting that the RELC should dedicate a major seminar to the subject of English Language Testing in its endeavour to improve standards of English teaching and learning in our member countries.

Thus the objectives of our seminar are:

1. To review and assess the practice, uses and techniques of current English language testing in South East Asia.
2. To discuss relevant new approaches and techniques in language testing and their application to the improvement of English programmes.

3. To provide the opportunity for the free exchange of ideas on the solution of problems in English language testing.
4. To stimulate critical thinking and efficacious activity among professional people concerned with English language testing in South East Asia and elsewhere.

The Seminar is planned in two parts. The first part will be devoted to the plenary session papers and discussion of these papers. A panel presentation of the objectives of English language teaching and the present status of English language testing in the SEAMEO member countries is planned for this afternoon. This panel presentation and discussion is scheduled in the first part of the seminar programme in order to help focus the seminar's attention on the South East Asian situation, its problems and its needs. The second part of the seminar beginning from Wednesday afternoon will be taken up by ten workshops each with a specific project to work on. The last plenary session, on Saturday, will be given to the presentation of workshop reports.

With such a wealth of scholarship and experience around us, the seminar promises to generate a good deal of new thinking and provide a lively and effective forum for the exchange of ideas and information. I express the sincere hope that everyone will find the week professionally profitable and stimulating and your stay in the beautiful city of Bangkok enjoyable.

Concurrently with the seminar, an exhibition of audio-visual teaching and testing materials has been organized with the cooperation of eighteen publishers and educational institutions. The RELC wishes to thank all of them sincerely for sending and setting up the large number of exhibits for the benefit of our participants. The display attempts to provide some information on the range and scope of materials available in the field of English teaching and testing. Our special guests are cordially invited to visit the exhibition after the session this morning. It is hoped that all participants will be able to find time to see the exhibits between working sessions in the course of this week.

May I now call upon the Director of the South East Asian Ministers of Education Secretariat, Gen. Netr Khemayodhin, to address us.

Address by Gen. NETR KHEMAYODHIN

Director, South East Asian Ministers of Education Secretariat

Your Excellency, Members of the Coordinating Committee, Mme. Director of RELC, Distinguished Delegates, Honoured Guests, Ladies and Gentlemen:

On behalf of the Secretariat of the South East Asian Ministers of Education Organization, I would like first of all to extend to each and every one of you a most warm and cordial welcome to the opening ceremony of this Fifth Regional Seminar organized by the Regional English Language Centre.

It is most heartening to see such a large gathering of participants, friends and well-wishers, and I hope that many of our guests and observers will be able to find the time and opportunity to attend the plenary and workshop sessions as well. This morning we are indeed fortunate and honored to have with us His Excellency Dr. Sukich Nimmanheminda, Minister of Education of Thailand, who has graciously consented to declare the seminar open. As Minister of Education of the host country, and as former director of the South East Asian Ministers of Education Secretariat, His Excellency Dr. Sukich has always had a deep concern for the welfare and progress of the South East Asian Ministers of Education Organization, its projects and activities. His presence here this morning further testifies to his continued interest, support and encouragement for SEAMEO and for what the organization is trying to do in and for the region in the field of education.

To date, as many of you may already know, the South East Asian Ministers of Education Organization has six regional centres located in the various member countries. They are the Regional Centre for Tropical Biology in Bogor, Indonesia; the Regional Project for Tropical Medicine and Public Health with a national centre in each member country and a central coordinating board located here in Bangkok; the Regional Centre for Innovation in Educational Technology, temporarily located in Singapore but ultimately to be located in Saigon; the Regional Centre for Science and Mathematics in Penang, Malaysia; the Regional Centre for graduate studies and research in Agriculture, in Los Banos, Philippines; and the Regional English Language Centre, in Singapore.

The Regional English Language Centre has for its primary objective the improvement of the standards of teaching and learning English as a second or foreign language in member countries of the South East Asian Ministers of Education Organization. In addition to its program in training, research, and instructional materials, the RELC attaches great importance to conducting regional seminars with the view to promoting the better understanding of English teaching and learning problems and to bringing together specialists in teaching English from member countries as well as from outside the region for concerted effort in finding solutions to the problems. To this end the RELC has, since it started its first operational year in July, 1968, conducted a number of seminars. For instance, in July, 1969, the Centre organized a seminar on "New Developments in the Theory and Methods of Teaching and Learning English". Although I did not have the pleasure of attending that seminar, at which some 200 participants from sixteen countries were present, I am nevertheless delighted to learn about its tremendous success. Apart from the teams of senior educationists and specialists in the teaching of English from the seven member countries of SEAMEO, there were also distinguished scholars in language and language education from Australia, Canada, Hong Kong, Japan, New Zealand, Republic of China, the United Kingdom and the United States.

For this seminar, the theme of which is English Language Testing, the planning committee aims to bring together some of the leading authorities on English language testing from both within and outside the South East Asian region as well as others concerned with the development, use and administration of tests and examinations. In this connection, it is gratifying to learn that the response to invitations has exceeded our expectations. Participants at this seminar will include members of the RELC professional staff, delegates from SEAMEO member countries, comprising leading educational specialists and top examination administrators as well as specialists in language testing from outside the region. With such a gathering of experts and qualified and interested participants, I have not the slightest doubt that the seminar will be a great success, and that its aims and purposes will be achieved. Judging from the program of the seminar it is evident that considerable time, thought and attention have been devoted to the planning, and I would like to take the opportunity to congratulate the RELC Director and the seminar planning committee on their fine efforts in the planning of the seminar.

To His Excellency the Minister of Education of Thailand I would like to once again express my personal gratitude as well as that of the Secretariat for hosting this seminar and for his kind presence here this morning to declare the seminar open.

To the Bangkok organizing committee which has rendered whole-hearted assistance and cooperation I also extend my appreciation for all they have done in contributing to the success of this undertaking.

While I am confident that participants will benefit from and long remember what will undoubtedly be an interesting, stimulating and fruitful session, it is my fervent hope that those of you who have come from other lands will take home with you many happy memories of your visit to this fascinating city of Bangkok.

Address by His Excellency Dr. SUKICH NIMMANHEMINDA
Minister for Education, Thailand

Mr. Director of SEAMES, Mme. Director of RELC, Distinguished Guests, Ladies and Gentlemen:

It is indeed an honor for this country to play host to two important meetings of the South East Asian Ministers of Education Organization within the same month — the meeting of the Council of the organization, and now this seminar.

A seminar on education and technical subjects reassures us that our organization will and can develop fully through interchange of knowledge and experience among experts and scholars in the related fields of specialization, and therefore it is very pleasant and gratifying for me to greet you all and to extend to you a warm welcome on behalf of His Majesty's Government.

As this is an expert seminar, there is no need for me to expatiate in great length on the importance and usefulness of the English language. But allow me to recall to you that once in this part of the world the study of the English language was a privilege enjoyed only among the elite of the country. Our early English teachers were recruited from early missionaries, and strange as it may seem, and perhaps coincidental, the first English schoolmarm who reached this country also came from Singapore. I mean the notorious Anna Leonowens. But now English is becoming more and more accessible to the masses, yet we often hear complaints that the teaching and learning of English in the schools and universities at present are not satisfactory. We may note further that the people in this region who are closely related in geography and culture but are different in their languages still find it difficult to communicate through the medium of English. If our effort in promoting the teaching of English is successful this language will serve as a means of overcoming this obstacle and thus help us cross this invisible language barrier.

I must on behalf of Thailand's Ministry of Education express my gratitude to the Regional English Language Centre, under the brilliant directorship of Mrs. Tai Yu-lin, for bringing to Thailand a large number of leading authorities on English language testing. A symposium of such renowned scholars I am confident will stimulate the interest of our English teachers and will inspire them towards new and innovative means of testing and evaluating their work.

I now have the honor to declare this seminar open, and trust that it will be a great success, and I thank all the participants in advance for your contribution.

Address by M. N. CARDIFF

British Council Representative

The British Government has given active support to the Regional English Language Centre since its foundation. Although the Centre has only been established a short time, it already provides an encouraging example of regional co-operation, and its success must give confidence to those concerned with more recently conceived SEAMEO projects.

The theme of the seminar is well chosen. The development of accurate and valid tests is an essential factor in language teaching at all levels. The presence of so many experts on the subject should ensure that an important contribution is made towards the perfection of testing techniques in the countries participating.

On behalf of the British Government I wish the Fifth Seminar of the Regional English Language Centre every success.

Address by JACK JUERGENS

Director, A.U.A. Language Centre, Bangkok

As an official representative of the United States Government and also a recent newcomer to Bangkok, it seems presumptuous of me to welcome participants to this Fifth Regional Seminar on English Language Testing.

Nevertheless, as Director of the American University Alumni Language Centre in Bangkok, I deem it an honor and privilege to greet all visiting participants and to welcome them to this marvelous metropolis.

Judging from the large attendance here, it appears that we should all take courage and apply renewed vigor to develop dependable criteria for evaluating the achievement of thousands of persons in Southeast Asia who seek to communicate with peoples of other lands through the English language.

In the last twenty-five years we have witnessed a burgeoning development of English into the world language of technology, cultural interaction, commerce, economic development, diplomacy, and essential international communication. Some observers have estimated that between 800 million and a billion individuals in the world either know this language as their native tongue, or can communicate in English as a second language.

If this estimate is reasonably accurate, then your efforts are of more than passing importance to a significant portion of the peoples of the world.

Let us all deliberate, cooperate, and communicate in order that we may depart from this seminar even better equipped to discharge our responsibilities. To this end we at A.U.A. pledge our unstinting support.

Again, welcome to Bangkok and the Fifth Regional Seminar on English Language Testing. May this week's labor prove fruitful.

Address by R. FINKELSTEIN

Representative of Canadian Government

Only three years ago my country celebrated its centennial. Blessed with prosperity, healthy, educated citizens and abundant natural resources, we are a fortunate land indeed. Most Canadians can look forward, eagerly, to an even better and more promising future. For many of the world's peoples the prospects are not as bright, the expectation in the years to come not as pleasing: for them, there is no hope in hunger; no gaiety in disease; no happiness in illiteracy; no charm in unending misery — and perhaps, no promise in life itself. Several centuries ago, an English poet wrote that "no man is an island sufficient unto himself" — civilization has moved so far and so fast since then, that today, we can say no city, nation, or even continent can long remain outside the forces of modern life — for all mankind lives in one world and it is indivisible. Because this is so, we must strive to provide the basic tools: the ability to read and write, to every human being. These skills will enable each individual to work, to provide for his family, and to sustain his dignity. We must give each person the chance, which will in turn allow the next generation to attain and achieve at an even higher level. There is almost nothing more commendable, or praiseworthy, than the tasks of you who are engaged in the educational field, for your work is basic to human wants and needs along with a plentiful supply of nourishing food, and adequate medical and public health services. On behalf of my government and the people of Canada, may I wish you well in your endeavours at this seminar and in the years to come. You are truly catalysts in economic development and social progress.

While Canada is generally considered a highly developed industrialized state, it is also in the ranks of developing countries, with vast reaches of its northern lands unexplored, untapped and almost unpeopled. Canadians are deeply and sincerely committed to international development assistance and anxious to contribute to this great work. A few weeks ago, my government announced the establishment of a unique international development research centre which will bring to bear modern technology and science on the problems of development, including those of education.

The peoples of the world want bread, not bullets; clothes, not military uniforms; housing, not hovels; education, training and jobs, not idleness and despair. And yes, they want cars, and transistor radios, television sets and luxury items. Through education, and a steadily rising standard of living, the material possessions will become more and more readily available even in remote corners of the earth. More importantly, because of your dedication, and that of your colleagues worldwide — through education, mankind may finally achieve its greatest desire and its ultimate blessing: peace.

Address by J. P. STOCKS

Hong Kong Government delegate

On behalf of the Government of Hong Kong, I have pleasure in offering this goodwill message to all delegates and observers at the Regional Seminar on English Language Testing. Once again my Government is honoured to have an opportunity to participate in a conference organized by the Regional English Language Centre on behalf of the South East Asian Ministers of Education Organization.

Although Hong Kong, because it is not an independent territory, is not a member of SEAMEO, we have followed its progress and development since its inception and have sent observers when invitations have been extended. Last year observers attended a SEAMEO conference in Djakarta and a RELC seminar in Singapore. Towards the end of the year my Government was privileged to return some of this hospitality by hosting the sixth RELC Co-ordinating Committee meeting held for the first time outside the SEAMEO member countries. This year an invitation has been extended and I am pleased to record, accepted for one of our officers to attend a four-month course at the RELC Singapore. It is essential that Hong Kong maintains that involvement, in the same atmosphere of co-operation and understanding, because of the interests and problems we have in common with member countries.

In the Report of the Co-ordinating Committee meeting held in Hong Kong, there was published Mr. Owen's paper on the preliminary plans for this seminar. At the end of that paper, Mr. Owen expressed on behalf of the planning committee the fervent hope that it would be an international seminar and requested any advice or assistance that could be given to achieve that goal. Judging by the large number present today from so many different countries — seventeen I have just learnt — the committee can feel the hope has been realized: it follows naturally that the success of the seminar is assured.

May I conclude, then, by expressing my personal pleasure to be here for the seminar and, on behalf of my Government, convey best wishes to all who attend.

Address by Mr. S. N. SAHA

Representative of the Indian delegation

It is very, very kind of the SEAMEO to extend an invitation to the Ministry of Education and Youth Services to send a delegation to the Regional Seminar on English Language Testing. The Ministry has gladly accepted the invitation and sent a two-man official delegation with Dr. Ramesh Mohan, Director, C.I.E., Hyderabad as its leader and myself of the N.C.E.R.T., an autonomous organisation under the Ministry of Education and Youth Services, as its member. I take this opportunity, on behalf of the Ministry of Education and on behalf of the leader of the team, Dr. Ramesh Mohan, of acknowledging with profound appreciation and gratitude the invitation of the SEAMEO.

It is a fact to reckon with that while methods of teaching have brought about great improvement in most areas of language instruction, English language testing has not comparably progressed to fulfil its function as an important part of the teaching-learning process. This seminar on English language testing, organised close on the heels of the Seminar on "New developments in the theory and methods of Teaching and Learning English," organised in June, 1969, is indeed a welcome step, as it has been rightly and imaginatively highlighted in the aims of the Seminar that new approaches and techniques on Language Testing should be marked by their application to the improvement of English Teaching.

It is heartening to observe that this year is being celebrated as an International Educational Year, and this Seminar significantly synchronizes with it. What could be a nobler tribute to this International Educational Year than seeing English language testing outgrow its limited purpose and set off in a new direction to become a richer and a more rewarding experience for the students. The significantly striking thing about this seminar also is that it is not just confined to the SEAMEO member countries. In effect, with specialists in teaching and testing from many other countries, it has put on an international complexion. It is ardently hoped that collective ingenuity and shared thought and resources of the specialist-delegates will be brought to bear on the successful accomplishment of the objectives that the organisers of this Seminar have envisioned.

ELEMENTS TO TEST IN ENGLISH AS A SECOND LANGUAGE

(A Testing Program for Elementary, Secondary and Collegiate Levels)

by BONIFACIO P. SIBAYAN

Introduction

My original intention was to write on the elements to test in English as a second language from a theoretical point of view. As the preparation of this paper progressed, however, it became quite clear that it might be more profitable if I described the language tests that we have been constructing at the Philippine Normal College, the program we are developing, and the things that we have aimed to test in this series of tests for elementary, secondary, and collegiate levels. It might also be fruitful if I discussed with you some of the notions that we have tried to work on and for us to critically evaluate some of these ideas.

Before I proceed to discuss these tests and ideas in some detail, I would like to make it clear that these tests have been developed in our graduate program in linguistics and language testing. We try to hit the proverbial two birds with one stone by training graduate students who come from various parts of the Philippines in the science and art of language test construction and administration and, in the process, we develop tests.

In his *Aspects of the Theory of Syntax*, in the section "Linguistic Theory and Language Learning" Chomsky writes the following very disturbing words on the study of language learning:

> . . . it seems clear that the present situation with regard to the study of language learning is essentially as follows. We have a certain amount of evidence about the character of the generative grammars that must be the "output" of an acquisition model for language. This evidence shows clearly that taxonomic views of linguistic structure are inadequate and that knowledge of grammatical structures cannot arise by application of step by step inductive operations (segmentation, classification, substitution procedures, filling of slots in frames, association, etc.) of any sort that have yet been developed within linguistics, psychology, or philosophy. Further empiricist speculations contribute nothing that even faintly suggests a way of overcoming the intrinsic limitations of the methods that have so far been proposed and elaborated. (Chomsky, 1965, p. 57)

If we were to take this view very seriously and pursue its implications to its logical conclusions, I am afraid that we may have to give up not only teaching and testing language the way we have been doing during the past two decades, but also the entire language teaching program.

However, such a suicidal intention need not be carried out. Even granting that Chomsky is right, we should go ahead and as generative grammar shows us the way, we will try to raise the quality of our teaching, testing materials and procedures.

Thus, because of the influence of structuralism in our methods and materials during the past two decades and because the influence of transformational-generative grammar as far as materials and methods of teaching has yet to be felt, most of the tests we have developed during the past several years are based on the language-as-habit theory and on the contributions of structural grammar. We have, however, constructed some tests which would seem to satisfy the notion of testing the competence of the learner rather than his performance in the second language. I shall take this up when I discuss two of the tests we constructed and used for six years in the Philippine Normal College in identifying college students who need remedial assistance in their spoken and written English.

1. Our tests and scheme of test development

To date we have constructed twenty-one language tests, three of which have been standardized. (See Appendix 1, p. 19) Of these, nine are achievement tests, three of which are in reading, one in writing, four in grammar, and one in listening comprehension. There is also a class of tests we call classification tests which are in a sense achievement and at the same time proficiency tests in some sort of way.

There are ten tests for the elementary school, four for the secondary schools, and seven for college. Six of those for college are for testing entering freshmen students — two in grammar, two in speaking, and two in reading. One is an achievement test in reading for senior college teacher education students.

The tests constructed for high schools are those for listening comprehension for entering freshmen, (which are either achievement or proficiency tests) a test in grammar for freshmen, and one for the third and fourth years in vocational schools.

If you will now look at Appendix 2 (p. 20), you will see the taxonomic chart for the elements of language we test. You will also note that we have developed tests in various aspects of the four skills involved in language learning — those of listening, speaking, and the allied skills of reading and writing.

2. Objectivity and tests in speaking or oral production

One of the most important principles the teacher or language tester is taught is that tests must be objective. Maybe we have carried this idea of objectivity too far without having gained any insights on language and language learning in the process. What is the purpose of objectivity? It might be pertinent to quote from Chomsky (1965, p. 20) on this point:

> The social and behavioral sciences provide ample evidence that objectivity can be pursued with little consequent gain in insight and understanding. On the other hand, a good case can be made for the view that the natural sciences have, by and large, sought objectivity primarily insofar as it is a tool for gaining insight (for providing phenomena that can suggest or test deeper explanatory hypotheses.)

I interpret the foregoing to mean that it is better to gain insight with a test that is not so objective than to have an objective test that simply produces objective results but does not lend itself to the gaining of insight. It is for this reason that in testing oral production, which is very difficult, we have developed a test that somehow is short on objectivity but is useful in having the examiner or teacher understand the examinee. This test (Lopez, 1968) has five parts: Part I — Sound Production, Part II — Answering Questions, Part III — Formulating Questions, Part IV — Giving Assignments or Directions and Part V — Oral Reading.

Let me take up Part II of the test including the set of instructions and criteria for scoring to illustrate what I mean by a test being non-objective but nonetheless reliable if properly handled by a competent examiner. Part II, Answering Questions, consists of the following:

1. Why does man need to eat? (Score for readiness and grammar)
2. Why do we pay taxes? (Score for readiness and grammar)
3. Why are trees useful to man? (Score for readiness and grammar)
4. How does man protect himself from diseases? (Score for coherence and relevance)
5. Why do we write letters? (Score for coherence and relevance)
6. What are monuments for? (Score for coherence and relevance)

In order to score the test, the set of instructions include the following instructions and criteria:

1. The four criteria discussed below will be used in evaluating the examiner's responses.
2. The four criteria are independent of each other, e.g., a response that is grammatically correct but irrelevant should be given credit for grammar but not for relevance.
3. Only two criteria will be observed at a time.
4. The examiner should evaluate the examinee's response on the basis of the criteria indicated in the answer sheet.
5. The meaning of the criteria:

    (*a*) *Readiness* refers to how long the examinee pauses before he starts reacting. (1) If it took the examinee quite a long time to answer, he should not be given credit for readiness. (2) If the examinee answers after some time but in a straightforward manner then he should be given credit for readiness. (3) The examinee who asks repetition of questions for the third time should not be given credit for readiness. (4) The examinee should be given credit for readiness if he readily answers the question with the short form as: Why do men eat? In order to live. The examiner should then encourage the student to use the complete sentence form by saying "Please use the complete sentence."

    (*b*) *Grammar* refers to the internal agreement of the parts of examinee's response. (1) If the examinee's response has acceptable grammatical structure then he should be given one point credit for that criterion. (2) If the examinee's response manifests disagreement in number, tenses, gender or other attributes of grammar then he should not be given credit.

    (*c*) *Coherence* refers to the manner of delivery. Does it come out as a piece or does he stammer and make false starts?

    (*d*) *Relevance* refers to the relatedness of the responses to the task set.

When results of this test were compared with two outside criteria: entrance test results and students' ratings in English 101 (Grammar and Composition), the computed coefficients of correlation were .71 and .69 respectively, both significant beyond the .01 level of confidence. The computed *r* of the whole test is .88.

3. Tests of integrated performance based on competence

I would like now to bring to your attention two tests in grammar for entering freshmen students in collegiate normal schools. My main interest in bringing these two tests to your attention is to indicate what we have done with them in a program of language teaching and testing. I would like to share with you some of the insights we have gained from the use of these tests.

During the last two decades, there has been a growing dissatisfaction on the part of supervisors of schools and a good segment of the public regarding the low quality of English, spoken and written, by many graduates from various normal schools. This situation was soon recognized by the members of the off-campus staff of the College which is the arm of the College that supervises the teaching experience of college seniors before they graduate. The chairman[1] of the unit requested the Language Study Center to help the staff develop a program to give deficient students some degree of competence before they went out to teach.

The first first group of seniors who needed assistance were identified during the second semester of the academic year 1962. A classification test for freshman college students in normal schools which was constructed by one of the graduate students (Corpuz, 1962) was administered to the seniors. The cut-off score in this test for successful work in freshmen college is thirty-seven. Sixty students who did not make the cut-off score were interviewed to determine what kind of remedial work they needed. They were then divided into two groups. A team of instructors and graduate assistants who acted as drill instructors, correctors of papers, etc., gave the students exercises in story telling, explaining processes, drills in pronunciation and grammar, and in composition writing three hours every Saturday for seventeen weeks. The students were also required to write 30 to 50 word diaries which were gone over with them by the graduate assistants. The results were encouraging. Tests showed that the students greatly improved in their spoken and written English. Practically all the students indicated that they benefited greatly from the program.

We used the Corpuz test during the 1963–1964 academic year and the first semester of 1964. While we were satisfied with it, we felt that it should be improved and if possible standardized using a wider population sample because it was constructed on a population of 100 freshmen who took the test in April 1961. By the second semester of the year 1964 the improved version "Achievement Test in English Grammar for Freshmen in Government Normal Schools" (Plaza, 1964) was completed. The validity, reliability, and tentative percentile norms of the new test were derived from 1,509 cases in four big normal colleges representing the most important (largest) linguistic groups all over the Philippines. The reliability coefficient of the test was .88. We used this test from the second semester of the year 1963–1964 to the second semester of 1969.

The test consists of two parts. Part I is composed of sixty-one statements with blanks requiring the use of adjectives, verbs, subordinate clause constructions, idioms, prepositions and conjunctions where the examinee has a choice from four suggested answers for each sentence; Part II is composed of thirty-nine items, each item being composed of four choices, the first three being statements in which nouns, pronouns, adverbs, capitalization, and punctuation are critical. The fourth choice is (D) *No mistakes*. The direction for Part II reads:

> Each item below contains three sentences. These sentences may be correct or one of them may have a mistake in grammar. If all of the sentences are correct, encircle (D) on your answer sheet. If one of the sentences is wrong, encircle the letter of that sentence on your answer sheet.

What is of interest to us is that the test seems to satisfy what Carroll writes about when he says that

> from a practical point of view it may often suffice to construct tests that measure only integrated performance based on competence. (That) . . . a general test of proficiency in a foreign language is often found to yield just as good validity when its items are complex, each drawing upon a wide sample of linguistic competences, as when each item has been contrived to tap competence in one and only one specific feature of the foreign language. (Carroll, in Davies, 1968, p. 56).

Our experience with the "remedial" students showed that if the students' competence was low as revealed by the tests, their performance would also be very poor[2]. This is what could have been suggested by Carroll (in Davies, 1968, p. 50) when he wrote,

> If 'competence' is a matter of whether a habit is present in the individual, 'performance' is partly a matter of the 'strength' of the habit — that is, the inferred degree to which the habit can be elicited, the rapidity of responses based upon it, and the extent to which it resists interference from other habits.

4. Tests in depth — tests that measure specific aspects of competence

We also have developed tests that measure specific aspects of competence or tests in depth. These are mainly for diagnostic purposes. Three of these are "A Diagnostic Test in English Prepositions for Grades Four, Five and Six," (Durian, 1965), "A Diagnostic Test on English Structures of Coordination for the Fourth, Fifth and Sixth Grades," (Olbes, 1966), and "A Diagnostic Test on Subject-Verb Agreement for Grades Four, Five and Six" (Parafina, 1966). The best of the three is that on coordination which covers forty-two structures involving *and* relationships. These are:

A. *Parallelism*

(1) Nouns in series, (2) adjective and adjective, (3) verb phrase and verb phrase, (4) adverb and adverb (place), (5) adverb and adverb, (6) infinitive and infinitive

B. *Single and Double Coordinate conjunctions*

(7) *and* (between subjects), (8) *as well as* (between subjects), (9) *or* (between subjects), (10) *and also* (subject and object position), (11) *either . . . or* (with more than a word between), (12) *neither . . . nor* (in subject and object positions),

C. *Reductions*

(13) *and so* + V aux + Nc, (14) *and neither* + V aux + Nc, (15) *but* + Nc + V aux *and* + Nc + V aux + *too*, (17) *and* + Nc + V aux + *either*

D. Sentence connectors that pattern with *and*:

(18) *and so* (to express result), (19) *but* (in contrast), (20) *thus* (as sequence signal), (21) *however* (as sequence signal), (22) *therefore* (as sequence signal)

Grammatical interpretations with use of

E. Double and single coordinate conjunctions:

(23) *both . . . and*, (24) *as well as*, (25) *neither . . . nor*, (26) *or* (in affirmative constructions), (27) *or* (in negative constructions)

F. *But* group

(28) *not only . . . but* (*also*) (both are true), (29) *not . . . but* (only one is true), (30) *but . . .* (ideas that contrast. Both are true), (31) *but* (as further reduction of *but not*) (*but* after verb phrase), (32) *but* (as further reduction of *but not*) (*but* after verb phrase)

G. *Reductions*

(33) *and* + Nc + V aux + *too* (addition of subjects), (34) *and* + Nc + *too* (addition of objections), (35) *and so* + V aux + Nc + Np

H. *Punctuation*

(36) *but* (reduction of *but not*), (37) *but* (as conjunction), (38) before *too* (mid-sentence), (39) words in series, (40) *so* before last word of the sentence, (41) before connectors, and (42) after connectors.

The test on prepositions (Durian) deals with fifty-one difficulties, namely

(1) prepositions used with transportation conveyances (*on* a bus, train, bicycle, public passenger plane; *in* a jeep, car, taxi, carretela, banca, a private plane; *off* a bus, train, scooter, a public passenger plane; *out of* a jeep, a car, a taxi, a carretela, a banca, a private plane.)

(2) Prepositions used with places

(3) Prepositions used with time

(4) Prepositions used with position and relationships

The test on subject-verb agreement (Parafina) deals with seven groups or sets of difficulties namely

(1) Pronoun as subject

(2) Problems with determiners in the subject

(3) Nouns as subject

(4) Problem of coordination in the subject

(5) Problems of inversion or inverted subject-verb agreement

(6) Problems of double agreement

(7) Problems with intervening materials between subject and verb

5. The triumvirate in teaching-learning activity

Allow me now to discuss a concept we are trying to develop with the use of tests as an integral part of the teaching-learning process. The universally accepted model in teaching looks like this:

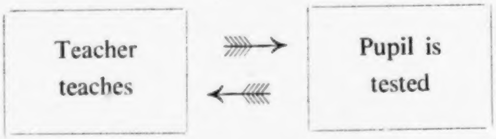


Fig. 1

The teacher teaches the child, a test is administered and the weaknesses of the child are discovered. The pupil then returns to the teacher for remedial teaching. This is the classical teacher-pupil relationship.

In the model we are trying to develop[3], the teacher teaches, the pupil is tested with a good test based on the course of study (syllabus is used by some), and instead of the pupil going back to the teacher he is given a set of self-instructional materials based on the test and on previous work for remedial work (Fig. 2). We have taken a test "A Standardized Achievement Test in Grammar for Grade IV" (Sibayan, 1966) as a basis for another graduate student to prepare self-instructional materials (Dedel, 1970) for this purpose.

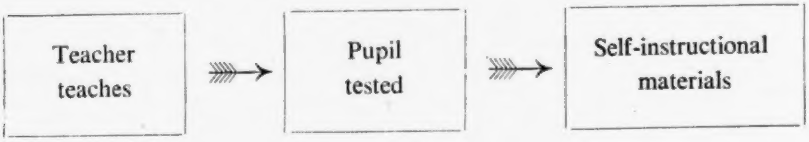


Fig. 2

While this model for the teaching-learning process is not really new, it is new in the sense that a standard test is used as basis for self-instructional materials.

6. Concluding Statement

Finally, I would like to say that although we have constructed these tests and continue to construct many more, the work is just beginning. At the same time we must try to train a group of teachers and supervisors in language testing and in the use of these tests in integrated programs.

Footnotes

1. Dr. Josefina Pulido, former chairman of the Off-Campus unit, presently Chairman of the Section on Professional Orientation, Graduate School, Philippine Normal College, Manila.

2. In order to gain greater and deeper insights into the language difficulties and causes of difficulties of the remedial students, one of our graduate students set out to study sixty-three remedial cases. The results of the study revealed that those who were identified by the Classification Test as needing remedial instruction compared to a group of non-remedial students selected at random from ten sections of seniors were students: (1) whose parents were less educated, (2) who came from schools with no or inadequate library facilities in both elementary and secondary schools, (3) who had much lower average ratings in high school, (4) performed worse than non-remedial students in their academic subjects in college, (5) did not have facilities for home study, (6) who almost had no hobbies. (See Biteng, Oralla. "The Intensive Course in English of the Philippine Normal College for the Schoolyear 1965–1966: Its Characteristic Features and Accomplishments." Unpublished M.A. thesis, Philippine Normal College, Manila, 1967).

3. This is being developed under the leadership of Dr. Edilberto P. Dagot, Coordinator of the Reading Center of the College.

REFERENCES

Books:

Carroll, John B. 1966. "The Contributions of Psychological Theory and Educational Research to the Teaching of Foreign Languages," in Valdman, Albert (ed.) *Trends in Language Teaching.*

Carroll, John B. 1968. "The Psychology of Language Testing," in Davies, Alan (ed.) *Language Testing Symposium.*

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass. The M.I.T. Press.

Davies, Alan. (ed.) 1968. *Language Testing Symposium. (A Psycholinguistic Approach*). London: Oxford University Press.

Harris, David P. 1969. *Testing English as a Second Language.* New York: McGraw-Hill Book Company.

Lado, Robert. 1961. *Language Testing.* London: Longmans.

Pimsleur, Paul. 1966. "Testing Foreign Language Learning." in Valdman, Albert (ed.) *Trends in Language Teaching.*

Valdman, Albert (ed.) 1966. *Trends in Language Teaching.* New York: McGraw-Hill Book Company.

Valette, Rebecca M. 1967. *Modern Language Testing.* New York: Harcourt, Brace & World, Inc.

Unpublished Materials:

(M.A. theses, Philippine Normal College, Manila)

Corpus, Monico. "Freshman Classification Test in Grammar for the Philippine Normal College." 1962.

Dedel, Xenia G. "Self-Instructional Grammar Materials for Grade Six." 1970.

Lopez, Lydia. "An Oral Production Test." 1968.

Olbes, Concepcion. "Diagnostic Test on English Structures of Coordination for the Fourth, Fifth and Sixth Grades." 1966.

Plaza, Maurita. "Achievement Test in English Grammar for Freshman in Government Normal Schools." 1964.

Sibayan, Felicidad. "A Standardized Achievement Test in Grammar for Grade Four." 1966.

Appendix 1

LANGUAGE TESTS CONSTRUCTED AT THE PHILIPPINE NORMAL COLLEGE

Author	Year Completed	Achievement	Classification	Diagnostic	Survey	Listening	Speaking	Reading	Writing	Grammar	Prepositions	Coordination	Elementary	Secondary	Collegiate
1. Alino ..	1966	×						×					6		
2. Carlota ..	1964		×							×				f	
3. Corpus ..	1962		×							×					f
4. Dominguez ..	1963		×				×								f
5. Durian ..	1965			×							×		456		
6. Gamelo ..	1966	×							×				5		
7. Guinid ..	1962		×					×							f
8. Habitan	1964	×						×					4		
9. Halili ..	1965	×				×								f	
10. Lopez ..	1968			×			×								f
11. Mejorada	1964		×			×								f	
12. Olbes ..	1966			×								×	456		
13. Parafina ..	1966			×						×			456		
14. Plaza ..	1964	×								×					f
15. Ramiro ..	1965	×								×			4		
16. Ramos ..	1967	×						×					4		
17. Reyes ..	1967			×		×		×					3		
18. Salva ..	1967				×			×							f
19. Sibayan ..	1966	×								×			4		
20. Sosa ..	1968				×			×							sen
21. Vergara ..	1966	×								×				j/sen v	

f – freshman year
j – junior year
sen – senior year
v – vocational

1. LISTENING

Areas tested	SUBJECTS TESTED														
	Elementary						Secondary					College			
	1	2	3	4	5	6	1	2	3	4	5	1	2	3	4
1. Sound Discrimination			×				×								
2. Word Discrimination							×								
3. Juncture			×												
4. Stress and Intonation							×								
5. Discourse Level															
(*a*) Speech Comprehension							×								
(*b*) Paragraph Comprehension			×												
2. SPEAKING	1	2	3	4	5	6	1	2	3	4	5	1	2	3	4
1. Sound Production												×			
2. Sound Discrimination												×			
3. Word Discrimination												×			
4. Stress and Intonation												×			
5. Oral Reading												×			
6. Discourse Level															
(*a*) Explanation												×			
(*b*) Formulating Information Questions												×			
(*c*) Giving Assignments												×			
3. READING	1	2	3	4	5	6	1	2	3	4	5	1	2	3	4
1. Letter Discrimination			×									×			
2. Vocabulary			×		×										
(*a*) Morphological Endings			×		×										
(*b*) Word Substitute					×										
(*c*) Syntactic Meaning					×										
3. Paragraph Comprehension			×		×							×			×
(*a*) Ability to recognize the literary devices used in the passage to determine its tone and mood.															×
(*b*) Ability to answer questions that are specifically answered in the usage															×
(*c*) Ability to select the main thought of the passage															×
(*d*) Comprehension and interpretation of informational material					×							×			
4. Interpreting Graphs, Maps, Locating Information					×										
5. General Reading			×									×			
6. Oral Reading							×								

4. COMPOSITION AND WRITING

Area of test	SUBJECTS TESTED														
	Elementary						Secondary				College				
GRAPHICS	1	2	3	4	5	6	1	2	3	4	1	2	3	4
Abbreviations				×										
Alphabetizing				×										
Capitalization														
Contractions				×										
Hyphenation											×			
Letter Format					×									
Punctuation				×	×	×	×				×			
GRAMMAR														
Adverb Order							×							
Agreement (Verb and Subject)				×	×	×	×				×			
Adjectives				×			×				×			
Antecedents						×								
Comparison Construction											×			
Conjunctions				×		×					×			
Case							×		×	×				
Determiners				×	×	×								
Grammatical Interpretation				×	×	×								
Grammatical Usage				×	×	×	×				×			
Gender							×				×			
Modifiers					×									
Modals									×	×				
Nouns and Pronouns				×	×	×					×			
Prepositions				×	×	×	×				×			
Tense						×					×			
Verbs				×			×				×			
Word Substitute				×										
Number							×		×	×				
SENTENCE AND DISCOURSE LEVEL														
Active and Passive Relationship						×								
Basic Patterns									×	×				
Changing Statements to Questions														
Direct and Indirect Discourse														
Sentence Sense											×			
Tag Questions and Rejoinders				×										
Thought Organization					×									
Transformations									×	×				
Immediate Constituents				×					×	×				
Inverted Subject				×	×	×					×			
Idiom											×			
PROCESS														
Dictation					×									
Following Directions, Using References, Graphs, Maps				×		×								
SPELLING														
Spelling				×	×		×							

Discussion

(*Discussants:* Sibayan; Bauer, Briere, Cherrier, Davies, Moller, Nababan, Saha, Thebus)

It is obvious, but in practice often overlooked, that testing and teaching are parts of one process, that testing should be based on teaching objectives, and that teaching should utilize the feedback furnished by testing. In this context the program described by Dr. Sibayan in which students are furnished self-instructional materials for use in re-studying items missed on a test is of considerable interest as a practical application of integration of testing and teaching. It also suggests a valuable area for materials development projects.

In this context, also, certain general problems appear. Objectivity is highly desirable in testing; to obtain objectivity language tests are organized on a taxonomic base, testing discrete items of language. Teaching also tends to be organized on the basis of a language taxonomy. Yet we all admit that the desired end-product of our teaching is communication competence, a global objective that is not defined by a language taxonomy. There is an obvious requirement for research into the various skills that constitute communication competence, and their interrelations. In the current state of the art, however, test developers have to fall back on taxonomic classifications, if only to furnish "labels" for discriminating and manipulating the data. There is not yet any description of communication competence that could serve adequately as a basis for test construction, or for language teaching itself. It would seem that development along taxonomic lines has reached an end-of-the-road stage; i.e., a stage where only minor further progress can be expected without some radical breakthrough to a new basis of development. Language testing is now perhaps in a stage comparable to the last developments of propellor aircraft, when jet propulsion furnished the breakthrough. It is not clear what the new propulsion system for language testing, and language teaching, might be. The current rapid advances in linguistic theory do not seem to have had much effect on language teaching.

The problem of oral testing brings some of these problems into sharp focus. Oral production is difficult to test, for many reasons. A competent oral examiner must be able to establish rapport with the student in the face-to-face testing situation, otherwise student response is inhibited and test results invalid. Many otherwise qualified people are not able to establish such rapport; in a group of 25 possible examiners of adequate language competence and background, one cannot expect to find more than three or four really good oral examiners. There is also the problem of the objectivity and standardization of oral testing. Yet one should beware of making a fetish of objectivity at the cost of insight. A good oral examiner can make valuable and accurate judgments of the communication effectiveness of the student, but these judgments are like those of an expert wine-tester — accurate, but often impossible to describe in "objective" categories. We cannot expect a high level of standardization of oral test results, especially if there are a large number of examiners involved, although careful training of oral examiner personnel can remove the grosser discrepancies. But if we lower somewhat our standards of objectivity, we will probably benefit from a gain in insight.

PANEL DISCUSSION ON OBJECTIVES OF ENGLISH TEACHING AND PRESENT STATUS OF ENGLISH LANGUAGE TESTING IN SEAMEO COUNTRIES

Chairman — PAUL A. SCHWARZ

TWO ASPECTS OF ENGLISH LANGUAGE INSTRUCTION IN INDONESIA

by P. TIRTOPRAMONO

I. The Objectives of English Language Instruction in Indonesia

It is a fact that the great majority of our university graduates today show a very poor command of the English language, a language by which they might keep themselves abreast of the scientific discoveries and advances made in their respective fields. Because of this, there appears to be a growing realization, among the academic circles in particular, of the need for a more rigorous and more effective English instruction as a means to implement the "maximum development of human and economic resources" of the country.

Although, historically speaking, the need to improve English instruction dates as far back as the early 1950's, due to unfavourable political and economic conditions the country has been able to do but a little more than 'patchwork'. This is not to say that not much has been accomplished. A great deal of good work has indeed been done in the right direction. However when it is viewed against the colossal scope of the country's problems — such as far-flung inter-insular communication, diverse regional conditions, and all the country's many gaps in keeping up with the pace of the modern world — the accomplishment does look meager. Especially so when we consider the fact that no positive results, unfortunately, have as yet been yielded by the efforts. Desired results wrought through educational measures always seem slow in coming into being.

Since 1952, various foreign aids have been involved to a great extent in the training and upgrading of teachers, the preparation and publication of lower secondary school materials, the carrying out of teachers' workshops, the present English Materials Development Project responsible for the preparation of higher secondary school materials as a continuation of the lower secondary materials, and the English Language Teachers Upgrading Project that presents the lower secondary materials and the techniques of handling these materials to English teachers of the lower secondary schools. In addition, there is also another project that provides a graduate scholarship programme for English teachers, at the Institute of Teachers Training and Education, at Malang. Referring to the lower secondary materials, it should be noted that since the use of these materials demands basic knowledge of linguistic principles that underlie the aural-oral approach, measures have also been taken to effect the inclusion of such a course in the curriculum of various teachers training courses. One principal move towards this end was the carrying out of a Seminar at Puntjak, in September last year, that was attended by some fifty-four delegates from the English departments of the Institutes of Teachers Training and Education throughout the country. However, since as far as the classroom instruction is concerned, much of this effort is preparatory rather than operational, it is hardly possible at this moment to expect any practical results. It is therefore more appropriate to consider these activities as indicative of the national awareness of the need not only to improve English instruction, but also to do it on a concerted, nation-wide scope.

Realizing that such a concerted effort presupposes at least certain uniform guidelines, the Puntjak Seminar discussed the following topics:

1. the teaching objectives and the curriculum;
2. the methodology and the course content;
3. the training and the upgrading of teachers.

As regards the objectives, the Seminar confirmed the ones as stipulated in Ministerial Decree No. 096/1967, which, in turn, was itself the product of a realization of the importance of accelerating the effort to improve English instruction throughout the nation. (The Decree, issued in December 1967, was to provide a basis for the execution of especially the English Teachers Upgrading Project, set up at about the same time, as an emergency programme. It must be remembered that education in general suffered seriously during the aftermath of the abortive coup of the Indonesian Communist Party in September 1965).

A stipulation of the objectives of the teaching of English in Indonesia (or in any other country for that matter), requires an understanding of:

1. the status of the language in the country;
2. the functions that it fulfils in the nation's life.

To learn something of the status and functions of English in Indonesia, we should take a glance into history and sociology. As early as the second half of the 1940's, that is, soon after the proclamation of the nation's independence, the country's educational leaders had already realized that the status of English in Indonesia was to be that of the first foreign language. The Indonesian coat of arms depicts the Indonesian community as one of unity in diversity. Indonesian, being the national language, is one symbol of such unity. Within this unity, which consists of diverse ethnic groups, each group has a dialect of its own. It is clear, therefore, that for most Indonesian children, Indonesian is their second language. They have to learn it at school before they can participate intelligently in the affairs of their country. But being a relatively new language, Indonesian has not developed to such an extent that it can replace the role of foreign languages, especially English, in the national pursuit and development of various branches of science and knowledge. In addition to this, the country has to adopt English as the international language, to carry out its international relations and policies. These are the factors that constitute the functions of English in the nation's life. As specifically stipulated in the Ministerial Decree No. 096/1967, these functions read as follows:

(*a*) *functions:* as a means to:

1. accelerate the process of the development of the nation and of the people;
2. establish friendly relations with other people;
3. conduct our foreign policies;

This stipulation provides the basis for the formulation of the objectives of the teaching of English in the country, which in the Decree are stated as follows:

(*b*) *objectives:* to lead to a "working knowledge of English" which, when stated in detail, reads as follows:

1. effective reading ability;
2. ability to understand spoken English;

3. writing ability;
4. speaking ability.

which will eventually equip the university students with the respective skills to:

1. comprehend the contents of textbooks and reference material in English which constitute 90 per cent of the total reference material;
2. understand the lectures of foreign teachers working under affiliation programmes, or communicate with foreigners including foreign students;
3. take notes on lectures delivered by foreign teachers, and introduce the Indonesian culture to other peoples;
4. enable them to make oral communication with teachers, individuals, and foreign students.

These are the objectives that constitute the groundwork of all the present activities directed towards the improvement of English instruction in Indonesia.

II. The Present Status of English Language Testing in Indonesia. The Problem of Scope

Conventional standards require that a good educational programme include an evaluation programme. In Indonesia, evaluation programmes constitute a subject of academic interest rather than a principle in operation. Therefore, a discussion of something about testing in Indonesia should be set not against the broader background of evaluation, but against the more narrow background of measurement. Educational measurement, as it is understood, employs techniques and instruments that produce the quantitative data of the observed properties. One such datum is the test score. The process of testing does not end with the assignment of these scores to the students. Interpretation must be made of these scores to find out a more accurate picture of the group of students taking the test, and how an individual student stands in relation to the rest of the group. Analysis must also be made of the items to see if the test really measures what it is supposed to measure. Very few teachers, however, get this far in making use of their test results. Most are content with obtaining the grades only. Also, due to community pressure against rigid grading, teachers tend to be lenient in their grades. It is even a common practice to change the norm so that more students may pass the examination. Together with so little use being made of the test results, the changing of the norm reduces the significance of the examination practically to nil. In the universities, however, the situation is better in some respects. For one thing, the more autonomous status of the university has to an extent shut off the influence of the community over grading.

III. Practices in Educational Measurement in General

Secondary school examinations in Indonesia are constructed by official committees. These committees are under the Directorate General of Education of the Department of Education and Culture, and may be organized at any of these three levels: local, provincial, or national. As regards English, a required subject in the secondary schools, up to 1967 the lower secondary school examinations were constructed at the national level; since then they have been constructed by the provincial committees. Though such 'transfer of responsibility' allows for more democratic examination as far as the individual school is concerned, these English final examinations are still external examinations. As such they suffer from the usual shortcomings of external examinations. Among these the most obvious is that very often they lack validity. It is a commonplace complaint among students after examinations that such and such items are beyond their level or scope, or that the examination, despite the fact that all the items are within the range of the students' achievement, is too long to be completed within the time alloted. Also, many examinations allow special factors to come into play to the disadvantage of many students. Such factors as speedy writing, exceptionally good memory and high intelligence, aural or visual defects, can often be easily detected in these examinations. In such cases, it is these factors, and not so much the student's achievement, that determine whether or not he passes a particular examination. Another shortcoming of an external examination that is all too common in our school system is the 'enthusiasm' with which teachers cram instead of teach their subjects, to meet the goals set by these examinations. In a situation like this, learning motivation has narrowed down simply to passing the examinations, a practice that undermines rather than promotes the ideals of education.

Procedures of testing and examining in Indonesia constitute so much a part of the educational setting that they have literally become mechanical, so mechanical indeed that as far as the school is concerned, they are part of the 'school ritual' year after year. Uses of test results are more administrative than pedagogical. Such uses seldom go beyond the assignment of grades for grouping, streaming, passing, or admission purposes. Admissions seldom take into account the student's past scholastic records. This produces a situation very common in Indonesia in which the student's future very much depends on his accomplishment in one single examination. As regards the tests made by the teachers (unlike in America where testing is an industry, in Indonesia teachers must make their own tests), a casual examination made of these tests will reveal how superficially most of these teachers are acquainted with the problem of test construction. Many even possess no more than a chance acquaintance with the field. To these teachers the principles of testing and measurement would be a whole new domain of knowledge that would shake the very foundation of their old beliefs. This is especially true of teachers who have never had pedagogical training at all, and whose educational background represents something of a hotch-potch of various other fields of knowledge. To further aggravate the situation, economic pressure had made it impossible for many of them to use their spare time to 'catch up with the demand of the profession.' But this is true of many other overworked teachers as well. Together they are no more that the automatons of education, from whom highly constructive education is least to be expected. Although many of these teachers still believe in the loftier ideals of education, they have no power to do anything but leave these ideals alone, an attitude reflecting defeatism, which is also manifest in their apparent lack of concern in the betterment of testing and examinations. But in connection with this it should be remembered that institutional education in Indonesia is mainly the responsibility of the state, so in so far as inadequacies are inherent in the system, they are far beyond the individual teacher's capacity to cope with.

IV. The Status of Testing

Being a required subject in the secondary schools, English occupies a position of importance in the final examinations of these schools. It is also one of the subjects included in entrance examinations, on both the higher secondary education and university levels. A failing mark on English will spell failure to the student concerned in the whole examination.

English tests and examinations emphasize language mastery. Content, such as literary and cultural knowledge, is primarily the concern of the English departments. In secondary schools, tests and examinations on language mastery are usually centred on structure, vocabulary and reading comprehension. In the past, secondary school examinations also included excerpts for translation from English to Indonesian. Composition is required only in final and entrance examinations to English departments of the Faculties of Letters and of the Institutes of Teachers Training and Education. These entrance examinations and examinations for scholarship abroad, especially to English speaking countries, also require an interview in English. These English departments and foreign scholarship programmes keep a rigid examination standard. However, despite this fact, these English departments always register the greatest number of students every year, when compared with other foreign language departments. The recent inflow of foreign investments has given a spurt to the English 'rush', which by itself had already held a leading position. Needless to say, these foreign companies require English interviews with the Indonesian applicants. All this should give us the idea that the need for English in Indonesia is very urgent and pragmatic in nature.

Unfortunately, like most tests and examinations on other subjects, those on English often promote goals that are not exactly the same as the one stipulated for the instruction of English. Serving as an example is the exclusion, in part in some respects, of oral examinations, especially in the lower secondary schools, where English instruction emphasizes more and more the oral aspect of the language. Not wishing, at this point, to discuss possibilities, I would like to treat the subject of the types of tests and examinations used in some detail.

At present the use of the types of written examinations shows a tendency towards more and more preference for the objective examinations. The types largely in use at present are the true-false, the modified true-false, and the multiple choice types. Fill-in is very much used in the secondary schools, while the multiple choice type is frequently used in the university entrance examinations, often in connection with an English-Indonesian translation intended to measure reading comprehension. Two main reasons account for the present tendency. First, a consideration of time has made it impossible to use the subjective examinations on the large masses of students entering schools and universities nowadays. Second, educational authorities have come to realize that various aspects are involved in the process of using a language. A consideration of this principle requires examination papers to be more comprehensive in order to obtain a more realistic assessment of the students' linguistic aptitudes. If a student fails in one aspect, this does not necessarily mean that his overall language command is poor. Objective tests enable the test administrators to measure these various linguistic skills in a reasonable time. However it must be remembered that as far as theory and practice are concerned, they may coincide in the labels only. I seriously suspect, in this connection, that the situation in Indonesia is a case in example. Therefore, I believe that a review of such tests and examinations is necessary so that any existing gap between theory and practice may be detected immediately and subsequent measures be taken accordingly.

OBJECTIVES OF ENGLISH TEACHING AND PRESENT STATUS OF ENGLISH LANGUAGE TESTING IN LAOS

by PINHKHAM SIMMALAVONG

I. Objectives of English Teaching in Historical Perspective

In the following paragraphs I propose to present the objectives of English teaching and the present status of English language testing in Laos in historical perspective. During the presentation I will spend some time talking about the place of the Lao, French, and English languages in the national system of education, the teaching of English, and its problems, before I proceed to deal with the main purpose of this panel.

Recently, there has been a point raised with regard to the aims of teaching English in Laos. Those who are familiar with our historical background and those who actually teach in our schools know that French is the first foreign language and the language of instruction in Lao secondary education. They are anxious to know why English, another foreign language, has been taught to Lao students along with French.

The answer is historical. English teaching is not traditional in Lao education, but inherited from France. When France came to Indochina in the 19th century, she also transplanted her system of education to her overseas colonies, including Laos, and so English was introduced. English has remained in the secondary school curriculum until today, for the Lao system of education is still very much patterned on the French system.

In French secondary schools, English is taught for reasons of utility whereas in the Lao school program it was studied as part of the *culture générale* or general education. Unlike French, which is the medium of instruction and the semi-official language in government and business affairs, English is taught as a foreign language beginning at 5ème (8th grade) in the Colleges (junior secondary schools, technical, and comprehensive schools), in the Ecoles Normales (teacher training schools), in the Lycées (secondary schools), in the Buddhist Institute, in the "Institut Royal de Droit et d'Administration (School of Law and Administration) and in the Ecole Supérieure de Pédagogie (E.S.P. or School of Education), except in the English Section of the E.S.P. where English is used as a medium of instruction.

In early times, English was taught by French teachers and recently, for the most part, by native speakers from Australia, New Zealand, the United Kingdom, and the United States of America. It was studied as an academic subject by a more passive method from a textbook called "L'Anglais Vivant" which was written on the linguistic and cultural basis for French students in France. However, at present, English is taught by a new and more enlightened method, namely, the aural-oral approach which has been found more successful than the traditional method that put a great emphasis on grammar and translation. Newer textbooks are being used now: *L'Anglais par l'Illustration, L'Anglais par l'Action,* and *L'Anglais par la Littérature, English for Today, English 900, Situational English,* and other books published by Longmans and Oxford University Press.

Today, English has a new meaning in Lao and Southeast Asian settings. It has gained in importance and is on an equal footing with French in a comprehensive school, a newly established school that was put into operation for the first time in 1967. It is only in this type of school that Lao, the national language, is used as a medium of instruction and French and English are studied as foreign languages. The establishment of this type of school is considered as an important and revolutionary event in the history of Lao education. Never before was Lao allowed to be used as a language of instruction in a secondary school.

Despite its growing importance, English teaching in Laos still has many problems. These were presented by teachers of English from all over the country at the National In-Service Seminar on Problems of Teaching English to Lao Students in Laos held at the Ecole Supérieure de Pédagogie in January of this year. The problems reported were: lack of organized courses, lack of suitable textbooks, shortage of reading materials (these are needed to supplement existing textbooks until materials specifically for Laos can be prepared), lack of time (especially in the Ecoles Normales where students have to take both academic and professional subjects), lack of motivation (in some schools), large classes (ranging from 30 to 50), uneven background, lack of teachers, lack of coordination of English programs in the country, dissatisfaction with the nature and organization of the English examination in the Brevet (a public examination given at the end of the first cycle of secondary education), and finally, lack of philosophy for teaching English.

The absence of a clear statement of the aims of teaching English has caused several of the problems just mentioned.

Finally, the English Seminar made a number of recommendations, the first of which was: "That the Ministry of Education issue a clear statement of the aims of teaching English in Laos." The following aims were suggested:

(*a*) Means for international and regional communication both for official and commercial purposes.

(*b*) Tool for higher education (foreign scholarships).

(*c*) Tool for technical education.

(*d*) Background for a better job.

(*e*) Necessity for studies in the English Section of the E.S.P.

It can be seen from the above that English has been taught in Laos for a number of reasons and in a variety of institutions. There is no question about the fact that there is an increasing demand to learn English in Laos, and the teachers are now asking for a clear statement of the aims and objectives. The secondary school system is in a state of transition, and new courses are to be prepared.

The Ministry of Education is considering the aims suggested by the English Seminar for teaching English in Laos, and if these are accepted either in entirety or in part, a clearer statement of objectives will emerge.

The aims stated suggest the following objectives: for purposes of international and regional communication, the skills of listening, speaking, and reading are of paramount importance, and in some cases, writing.

For purposes of higher education in an English department or university, hearing, speaking, reading and writing are all of great importance.

For purposes of technical education and commercial endeavours, writing is less important than the other three skills.

It was not considered that the study of English literature was an important aim, but rather students should be given plenty of opportunity to read modern English prose and the kind of writing found in newspapers, magazines, and everyday communication.

The objectives for the Lycées and the Fa Ngum Comprehensive Schools would include all the four language skills mentioned above, and in the technical schools and adult classes, hearing, speaking, and reading would be of greatest importance.

To sum up, the objectives in teaching English in Laos appear to have moved from teaching it for the sake of literature and general culture to teaching it for more utilitarian reasons, for communication, and as a tool for further education.

II. Present Status of English Language Testing

As there is no national English syllabus, there is no national programme of English examination in Laos. The present English examination system is both internal and external.

At the Lycées both systems are in operation. *Internal* examinations are used at all levels during the school terms, but *external* examinations are used in public examinations which are given at the end of the school year, normally in June. The public examinations are the Brevet (short name for Brevet d'Etudes du Premier Cycle or B.E.P.C.), the Diplome (short name for Diplome d'Etudes du Premier Cycle Secondaire de l'Enseignement National or D.E.P.C.E.N.), the National Baccalauréat, and the French Baccalauréat.

Only *internal* examinations are used at the Ecoles Normales (schools for training primary school teachers), at the comprehensive schools, at the Ecole Supérieure de Pédagogie (school for training secondary school teachers), and at other institutions as well.

Since there are several types of schools in Laos that carry out English teaching programs, I should like to limit my discussion to the types of English tests and examinations that are used only in some schools at the following levels:

> *Brevet or Diplome:* this examination is given at the end of the first cycle of secondary education period. Formerly, a written and oral English examination was given in the Brevet or Diplome, and reading comprehension and translation (English-French) were used. In recent years, the examination has had only the oral part. However, the nature of the oral examination remains the same as it was before, i.e. translation and reading comprehension are again used. In the reading comprehension test, students are given a passage of about twenty lines to read silently for a few minutes then to read aloud at normal speed with appropriate expression. At the end of the reading, questions such as the following are asked: "What does this word mean? What's the general idea of the story?" and so on. The reading comprehension test is then followed by translation. Usually, sentences in the reading passage are given for translation. Each student gets 15 minutes to do the whole oral examination. In the reading comprehension part, each student is given a different passage that he has never seen before to read and is given different questions to answer. While one student is being tested in front of the room, the other reads silently in the back of the room.
>
> Although translation is of some usefulness, it is a very difficult art. Its purpose at this level, where students have just passed beyond the aural-oral stage of learning, is not very clear to many teachers. The reading comprehension test is usually used to measure a complex of abilities, i.e., the

abilities to understand the language, the ideas, and the style used in the reading passage. However, conducted in such a manner as described above, this test may not test what it wants to test.

Baccalauréat: this examination is given at the end of the second cycle (three-year program). The status of English in the Baccalauréat examination is the same as in the Brevet, i.e., it is oral and is not compulsory. The nature of this examination is also the same as that of the Brevet, but at a higher level. The types of tests that are used in the Brevet are also used in the Baccalauréat and for the same purposes.

Entrance Examination to the E.S.P.: this is a public examination to be given to candidates who wish to enter the English Section of the E.S.P. Beginning next school year, those who want to come in, after three years of English study in a secondary school, must take an entrance examination in English. This consists of two parts: written and oral. The first part includes several sub-tests on grammatical structure, vocabulary, reading comprehension, and composition. The second part includes sub-tests on aural recognition, aural comprehension, and oral production.

The purpose of this examination is to assess language aptitude and to measure achievement in written and oral expression.

Final Examination of the E.S.P.: this is an internal examination given at at the end of seven years of study in the English Section. It is composed of two parts: written and oral. The oral is made up of reading, seen and unseen, dictation for word endings, aural recognition, and speech, prepared and impromptu. The written consists of composition, (about two pages on one of several topics given) comprehension, and vocabulary.

The comprehension sub-test contains a reading passage of about 40 lines and some questions on the passage. In the vocabulary sub-test, several exercises are used: matching words, using words in sentences, answering questions in sentences, giving meaning of words in sentences, and the like.

The whole examination given is intended to measure achievement in written and oral expression.

Final Examination at the End of Intensive Course of the E.S.P.: This is again an internal examination given at the end of the intensive course of the first year English Section. This is designed to give students who have come in from primary schools, an intensive study (about twenty-six hours a week) of pronunciation, grammar-structure, and reading and vocabulary, so that they will be able to achieve the four language skills and will be able to study other subject matter in English when they advance to higher grades. At the end of the intensive course, a final examination is made up to measure the following skills or areas: aural recognition, aural comprehension, oral production, reading comprehension, vocabulary, grammatical structure, and written composition. Various item types are used.

Purpose	*Item Types*
Aural recognition	Words in isolation and in short sentences. (*a*) Discriminating whether there is a difference. (*b*) Indicating on answer sheet which sentence has been heard. (Two sentences given)

Purpose	*Item Types*
Aural Comprehension	(*a*) Listening to a story. Answering questions about it in sentences, or indicating TRUE or FALSE about statement on answer sheet.
	(*b*) Listening to a sentence or sentences and choosing one of four sentences on answer sheet which is true in relation to the story.
Oral Production	(*a*) Reading aloud words and phrases containing points of special difficulty.
	(*b*) Reading aloud a simple passage.
	(*c*) Answering simple conversation type questions.
Reading Comprehension	(*a*) Reading a story and answering questions about it in sentences. (Written expression is examined, also.)
	(*b*) Reading a story and then answering questions by choosing the correct answer of four given. (Multiple-choice or True/False).
Vocabulary	(*a*) Matching words with meanings.
	(*b*) Writing opposites to words given in context.
	(*c*) Sentence completion with the correct word — choosing words from list given.
	(*d*) Paraphrase (multiple-choice). Choosing one of four words or phrases that mean the same as the word underlined in a sentence.
Grammatical Structure	(*a*) Completion (multiple-choice)
	(*b*) Completion — supply correct form of verb pronoun, etc.
	(*c*) Writing Yes and No answers, long and short.
	(*d*) Reading sentences or short passages using special grammar points, then answering or completing answers to questions. The answers include grammar points being tested.
	(*e*) Conversion: changing to interrogative or negative.
	(*f*) Conversion: making questions to which answers are given. (Sometimes first word given)
Written Composition	(*a*) Writing sentences or short compositions on set topics. Pictures and/or words and guidance given.

The purpose of the final examination at the end of the intensive course is to assess achievement in various language skills presented above.

English testing has been listed as one of the many problems of teaching English in Laos. One of the main problems is concerned with the nature and organization of the English examination in the Brevet or Diplome.

Most English teachers are not satisfied with the Brevet or Diplome exam, and are of the opinion that the English examination should be conducted entirely in English without translation. They feel that translation is too difficult to achieve at this level where the student's English proficiency is still in its intermediate stage. Besides, students have to translate from English into French, another foreign language, which makes it twice or three times more difficult than translation in its usual sense and also means that the student's French is being tested as much as his English.

The teachers also feel that it will be more satisfactory from the teaching point of view if the examination is also made compulsory because English study is compulsory. However, this may not be possible, for not all students taking the Brevet or Diplome have had a chance to study English due to the lack of teachers. This situation is particularly true of small Colleges in provincial areas.

Apart from this specific problem in testing, there are the more general ones that arise from lack of clearly defined objectives and lack of enough qualified staff. If a truly national system of tests is to be adopted, there is a great need for someone to be appointed by the Ministry of Education with the responsibilities for co-ordinating the English programs in the country, distributing textbooks and reading materials, organizing examinations, and working as a liaison between the schools and the Ministry.

The first step has been taken to solve some of these problems. The Ministry fo Education has already given approval for a seminar to be held sometime this summer vacation to work out a syllabus for the first three years of English, leading to the Brevet or Diplome examination. At that time, it is hoped that the Ministry will be able to give a clear statement of the aims and objectives for teaching English in Laos, so that the syllabus can be worked out on these aims and objectives.

This will be the first national syllabus set by the Ministry of Education. There is a great need for this to provide the basis for the examinations, for the latter are linked very closely to the syllabus and these too are also linked with the aims and objectives.

OBJECTIVES OF ENGLISH LANGUAGE TEACHING AND PRESENT STATUS OF ENGLISH LANGUAGE TESTING IN MALAYSIA

by MURAD BIN MOHD. NOR

The testing of English Language is part of the national examination system in Malaysia. The local examinations other than the Overseas School Certificate Examination conducted by Cambridge, which all school pupils have to face, are the following:

1. Assessment Examination (taken at the end of the 5th year of primary education).
2. Standardised Examination (taken at the end of the 6th year of primary education).
3. Lower Certificate of Education Examination (taken at the end of the 3rd year of secondary education or 9th grade).

It is worthwhile to look into the purpose of these examinations before taking a close view at the objectives of English language testing.

The Assessment Examination was introduced in 1967 after the abolition of the Malayan Secondary Schools Entrance Examinations three years earlier. There were general complaints that a good proportion of the pupils were automatically promoted into secondary school without a sufficiently strong foundation. Although the complaints were quite unfounded, as it is obvious that the better quality of the pupils previously promoted into secondary school were the 60 per cent who were able to find places in secondary school after passing the Malayan Secondary Schools Entrance Examinations, it was decided that measures should be taken to ensure that pupils and schools know their levels of achievement, and, if necessary, take steps to bring up their achievement levels. It is hoped that remedial measures take place in Std. 6, the year before the pupils go to secondary school. As an insurance, another check test, the Standardised Examination, is taken at the end of the 6th year. Pupils automatically proceed into secondary school and after 3 years, they take the Lower Certificate of Education Examination. This is a certificate examination and a certificate is awarded for success in a required number of subjects. However in practice, the Lower Certificate of Education Examination is a selection examination and only those who acquire a Grade I certificate may proceed to upper secondary school. After 2 years, these pupils sit for the Overseas School Certificate Examination conducted by the Cambridge Examinations Syndicate. We are happy to inform the Seminar that since last year, the Malaysia Examinations Syndicate has introduced a new Examination — the Malaysia Vocational Certificate of Education which is specifically meant, as the name suggests, for the non-academic.

Without doubt, all the examinations the pupils face in their school career are achievement tests. Coming specifically to English Language, all papers at the various levels test ability to use the language.

English language teaching and testing are so geared to fall in line with requirements of the Malaysia Certificate of Education (Sijil Pelajaran Malaysia). As spelled out in the Examination in one paper, candidates are required to display in their writing the following: arrangement, subject-matter, general expression and command of the language. In the other paper, given a passage or passages of prose on which questions are set, candidates are tested on their ability to understand the content and argument of the given text and to infer information and meaning from it. Questions also set out to test ability to summarise. An oral examination consisting of reading and understanding a prose passage and a short conversation, looks mainly for pronunciation, intonation and fluency. However, failure in the oral examination will not prejudice the written result, although satisfactory work will be taken into account.

The various levels of English tests conducted by the Examinations Syndicate aim at testing those skills which ultimately converge to those spelled out for the Malaysia Certificate of Education (Sijil Pelajaran Malaysia) Examination. What those skills are, are guided by the English Syllabus for the Primary and Secondary Schools. The broad aim of the Syllabus is "to give pupils, through the skills of writing, listening with understanding, reading and speaking — a complete mastery of the essential structures of the English Language and of additional vocabulary and idioms." To achieve this aim, the syllabus lists for each grade level, the structural and linguistic items, reading material and exercises for oral and written expression. There being no specific skills defined in the syllabus for each grade level, the tests conducted by the Syndicate aim to test some skills laid out arbitrarily by someone and modified year after year based on opinions of individual setters of papers and those of panels whose members do not necessarily sit on panels for successive years.

Language testing, so far conducted by the Examinations Syndicate, cannot claim to have covered all the four major areas of language testing, i.e. listening, speaking, reading and writing. All the tests so far are pure paper and pencil tests. In the hope that the listening and speaking aspects are well looked after by the teachers, the language papers test pupils' achievement in reading and writing.

Let me describe generally the various types of English tests currently used in the different national examinations and look into the background for such tests.

Assessment Examination

1. *English I:* This is essentially a second language test. This paper is taken by pupils from the non-English medium schools, i.e. pupils who have their medium of instruction in Malay, Chinese or Tamil. Apart from the fact that the linguistic backgrounds of these three groups are different, English language teaching and learning also take different roles in the curriculum. The Malay medium pupil learns English as a second language from the first year of school, spending 180 minutes a week for the first three years, and 200 minutes a week for the next two years before he takes this English I test. The Chinese and Tamil medium pupils, on the other hand, learn English as a 'third' language and are introduced to it only in the third year of school; they have begun learning Malay as a second language from the first year. In the learning of English, the Chinese and Tamil pupils spend 120 minutes a week while in Standard 3, 160 minutes a week in Standard 4, and 200 minutes a week in Standard 5, and they are expected to come to the same level of achievement as their counterparts in the Malay-medium school. This has been spelled out in the English Syllabus for Primary Schools and hence only one English paper is offered to all these three media schools.

Thus from the very start, there is the problem of designing a paper for different groups of candidates. The paper attempts to cater for two groups of candidates, (two groups in terms of time spent in learning English) and is designed on the following lines. On a one-hour, 50-item multiple-choice (4 options) test, 20 items are devoted to testing reading comprehension and the remaining 30 items to testing writing within limitations.

In the reading comprehension test, candidates are given four passages of graded difficulty, the length of the passages varying between 100 to 200 words. The materials are presented in the form of a narrative, conversation or letter and five items are set on each passage.

The items test the following skills:

1. Ability to give a suitable title to the passage, or the main idea of the passage.
2. Ability to give the meaning of a selected word or words in the passage.
3. Ability to translate ideas; i.e. answer questions based on the passage, the answers to which are given in words different from the text.

In the writing test, candidates are tested on the basic grammatical structures classified under sub-topics like tenses, subject-verb agreement, prepositions, pronouns, etc. Candidates are not given a writing test to test ability to express, the multiple-choice test being the best we can offer at the moment.

2. *English II:* This is a higher level paper than English I and is meant to test those whose medium of instruction is English. The candidates who take this paper can be broadly categorised into two. On the one hand there is the group where English dominates the life of the pupil both at home and in school; at least one parent, usually the father, speaks English; the student meets in his non-school environment and understands and speaks English although there is every likelihood that often he hears and uses poor English. However you can be sure he has opportunities to read books in English. Very soon after a few years he even thinks in English. On the other hand, there is the group where the pupil hears and speaks English only in school. He lives in a community where English is hardly used. Neither parent speaks English and among his playmates even in school he is more at home conversing in his mother tongue. This group usually come from an economically poor community.

English II is also a one-hour, 50-item test built on similar lines. Part of the paper is a reading-comprehension test consisting of 25 four-option items. Candidates are given four graded passages each of approximately 120–150 words in length. Reading material consists of a narrative, a conversation, information or explanation and directions or announcements. The skills tested are:

1. Ability to translate ideas.
2. Ability to give the meanings of words.
3. Ability to give the main ideas or suitable titles.
4. Ability to give the sequence of events or actions.
5. Ability to make inference.

In the writing test, besides testing the conventional grammatical structures, some items are included to test organisation and sequence.

Lower Certificate of Education

To date, English Language Testing for this examination has been in the so-called traditional form. As from 1971, the English Language Test shall consist of two parts — a writing test and a reading comprehension test in the multiple-choice form. There are two English Tests at the Lower Certificate of Education Examination — English I being designed for Malay Medium candidates and English II for English Medium candidates. As in the Assessment Examination certain factors have to be borne in mind in the design and construction of the English II paper. Candidates taking English II consist of two groups — one which has had English as the medium of instruction throughout the primary school and another which has had either Malay, Chinese or Tamil as medium of instruction in the primary school but pursued their secondary education with English as the medium of instruction. This second group has been given a special one-year programme to learn English before entering the first year of secondary school to enable them to go through an English-medium education. English I is a two-hour paper consisting of: (*a*) a writing test, in which candidates are to choose the subject from a number of topics and write a composition of about 120 words in about 1 hour; (*b*) a comprehension test, in which candidates are to write out answers to questions based on the passage; and (*c*) English structure test items which usually offer a selection of choices or in the case of a test on verbs the infinitive form is given and candidates are to write out the current form of the verbs.

English II is a 2½ hour paper and follows the same pattern as English I; candidates however are to write a composition of about 300 words in about 1 hour.

Each of these two papers is set by one individual and the questions and marking schemes are reviewed and revised, if necessary, by a panel of which this individual is not a member. The chief examiner, who is neither the test-setter nor a member of the panel has the final say over the marking scheme. An officer of the Examinations Syndicate who sits on the panel meeting acts merely as a secretary. The unhealthy procedure of the setter, panel and chief examiner working independently in the building of the same paper, will be eliminated when the English papers are objectivised. The paper will be set by a panel; the chief examiner will be a member of the panel which will set the entire paper and draw up the marking scheme for the essay. The papers will take the following pattern:

English I — Part A — 1¼ hours consisting of (*a*) a compulsory controlled essay where marking is based wholly on English; (*b*) an essay on one topic selected from a choice where marks will be awarded for ideas, arrangement, vocabulary and English. Part B — multiple choice paper — 1¼ hours — 20 items on Comprehension based on 2 passages testing:

1. Ability to translate ideas.
2. Ability to give the main ideas, or title of the passage or part of it.
3. Ability to give meanings of words.
4. Ability to give sequence.
5. Ability to make inference.
6. Ability to analyse or synthesize or draw conclusions and 40 items to test structures.

English II will be designed on the same lines as English I.

Malaysian Vocational Certificate of Education

The English Paper requires candidates to show: (*a*) ability in writing and a command of the language, arrangement, general impression and subject matter in that order as specifically looked for; and (*b*) reading-comprehension. This part of the paper seems to demand skills parallel to that in the Malaysia Certificate of

Education (Sijil Pelajaran Malaysia) Examination. The testing technique is different, comprehension being tested through the multiple-choice type of questions. Candidates are tested mainly on their ability to translate ideas either by answering questions based on the given texts by selecting the correct answer from a number of choices given in language different from that of the text or by giving the meaning as in the text of words or phrases taken from the text. Candidates are also tested on their ability to synthesize or analyse; e.g. to give the theme or main idea of a paragraph or the whole text, or to give a suitable title to the text. This is answered by selecting the correct answer from a number of choices. Where applicable, inference and sequence tests are also given.

In addition to what has already been called for, in the writing test, part of the paper also tests English grammar. The questions set are first of the free-response though controlled type but in the second year, the pattern changes to multiple choice.

It is to be noted that the English test for MVCE Examinations is now designed in the objective form except that the essay writing component still remains.

Problems have always existed in the testing of English. The aims of English learning have only been generally stated. The designing of tests to test these aims has so far been left, in the case of traditional papers, to individual setters, the panel often not doing very much except to rewrite or replace glaringly poor questions. Often the setter is unaware of changes that may have been made. This is because papers are set about 2 years ahead of time. The setter is given the previous year's paper as a basis and is quite unaware of the pattern for the current year. The panel is also in the same plight. Hence it is not uncommon that one year's panel may consider a particular skill or area suitable while the next year's may not and the following year's panel may introduce it again. Hence there is great difficulty in developing a trend or some form of consistency in the tests.

Grading of essays always poses its problems. The method of assessing essays has moved from a somewhat controlled system of awarding marks under different skills like subject matter, sequence and language to one of marking by impression. With about 400 examiners and an essay graded only once, reliability is bound to be low. The shift is now back to the orthodox method of awarding marks according to definite skills. This, in a way, is brought about with the objectivising of the other sections of the Language Test and the present problems of affording at least two readers for each essay. Marking by the orthodox system also poses its problems.

The comprehension test has been questioned only recently. The practice is to provide a passage of about 250 words, and candidates answer about 5 to 6 questions to earn about 25 per cent of the possible score for the Language Test. Candidates are strictly instructed to answer in their own words and this instruction has often penalised those who attempted to do so, the reason being that the language of the passage is already so simply written that over-simplification leads to awkward sentence constructions. Besides, I wonder whether sufficient consideration has been given as to whether language in the comprehension answers should be marked down as heavily as 50 per cent of the possible score.

Often the comprehension tests do not test anything more than mere translation of ideas.

Some of these problems may be solved or at least there should be an improvement over the present situation when the English Tests are in the objective form.

Often again, the setter and the panel members moderating the paper do not have any knowledge of the performance of the candidates in the previous years' papers to guide them in setting the standards. This problem does not exist with the English papers in the Assessment Examination which are objective type. Both pre-test and post-test data are available to the test developer building the paper. The problem will be eliminated with the English papers objectivised at the Lower Certificate of Education level. However, reliability will be lower in the composition component of the paper. Marking time, running costs and the lack of examiners, do not allow multiple reading of the composition test to help limit wide variability in the scores.

To achieve greater reliability in the composition test, it has been decided that candidates are to be tested in two compositions — one a compulsory and controlled essay where language will be the only consideration for award of scores. In the other essay, the candidates will select a topic from a few choices and the composition will be marked for content and ideas, arrangement or sequence, vocabulary and grammar. The weightage of marks for each of these components takes into account the expected performance of an average candidate.

One main problem will always remain with testing until the syllabus spells out clearly the terminal behaviour at each standard and form. With only the general aims for the whole English syllabus, the task of determining the terminal behaviour rests mainly with the test developer who is continuously seeking the advice of his panel to determine the objectives and designing tests to test those objectives.

ENGLISH TEACHING AND TESTING IN THE PHILIPPINES

by CLODOALDO H. LEOCADIO

Allow me to extend to you all the cordial greetings of the teachers from the Republic of the Philippines. It is indeed a rare privilege to participate even in a minor way and share with you our limited experience in the area of English Language Teaching and Testing. I wish that even for a fleeting moment, since I realize that time is of essence, you would allow me to express the gratitude of the members of our delegation for the wonderful hospitality of the Thai teachers. We also would like to express our thanks to the members of the secretariat, to Dr. Schwarz, Mrs. Tai Yu-lin, and Mr. Owen, whose enthusiasm could readily be reflected in the various correspondence we received. To all these we give our sincerest appreciation, and congratulations for a very smooth coordination and management in the preliminary planning stages of this Seminar.

English in the Philippines is almost as old as this century now entering its seventies. From 1900 to 1957 English was the sole medium of instruction in the Philippine Public Schools. A whole generation of Filipinos educated in English speak and write it and use it in the conduct of their daily affairs. There is a remarkable body of literature written in English by Filipinos. Radio, television, the movies, and almost all metropolitan dailies in Manila use English. Business establishments, local and foreign, require a good command of English in applicants for jobs in their companies. Yet in spite of its age and continued prestige, English in the Philippines is becoming more and more difficult to teach. The problems of language teaching have plagued us for many years.

English teachers in the Philippines like to point to the present language policy in our public schools as the cause of their troubles. In our present set-up, the native vernacular dialects are used as medium of instruction in Grades I and II, with English and Pilipino (the national standard language) taught as separate subjects. Beginning in Grade III, English becomes the medium of instruction, with Pilipino taught as a subject through high school. In the colleges and universities twelve (12) units of Spanish are required for graduation. It used to be twenty-four. The English teachers' blame on such a set up for the increasing difficulty of English teaching seems an oversimplification of their manifold instructional problems. However, it is quite true that this multi-lingual situation has given rise to such complex problems as the training of teachers for effective teaching in three languages, the preparation of sound instructional materials for the first- and second-language learners of Pilipino and for second-language learners of English, and the evaluation of instruction in all three languages — Pilipino, English, and the local vernacular. There seems to be no simple solution in sight, but whatever the solution will be, English will continue to be used and taught as a second language in the Philippines because it is our "linguistic bridge to the outside world".

How successful has English teaching been in the Philippines? A generation of English-speaking Filipinos, engaged in various fields of endeavor, public and private, many of them occupying positions of leadership, all of them directly or indirectly involved in the task of nation-building, might be presented as the strongest tangible proof of successful teaching of English. On the other hand, a thoughtful few of the same generation could perhaps point to the millions of school children who drop out of school, profiting little from instruction given in a borrowed tongue. These dropouts would be mute, since they cannot be eloquent, *proof* that somehow English instruction has failed to touch the lives of the people in the small and far-flung villages and hamlets of the Philippines.

Such off-hand evaluation of English teaching, however, is wholely impressionistic and not based on actual figures. To evaluate the results of English teaching better we have to go back to the objectives of English instruction in the Philippines to determine what is expected of children learning English and how these expectations are to be evaluated.

The following objectives were stated by Dr. Antonio Isidro, retired President of Mindanao State University, in a work conference of supervisors and teachers of English in April 1953:

1. To make English an effective medium of instruction or a tool for understanding all other content subjects.
2. To cultivate it as a useful means of communication and of understanding Philippine society.
3. To use it to promote understanding and appreciation of other peoples and their cultures.

The attainment of the first objective is the concern of the elementary schools. It is during the period from Grade I through Grade VI that pupils are supposed to develop gradual mastery of the basic patterns of English to enable them to learn the content of other subject fields. It is assumed, therefore, that after two years of purely oral English in Grade I and II and with sixty (60) minutes daily of formal instruction in language and reading from Grade III through VI, besides the "practice" in the use of English that pupils get in the learning of other subject fields, Filipino pupils will develop enough fluency in the language and a fair degree of reading skill to enable them to tackle high school work.

Much indeed is expected of the Filipino learner of English even while he is still learning the language itself. He must use it not only to gain knowledge but to communicate such knowledge to his teacher and classmates. He must listen and read with understanding to learn about life around him, and later he must learn to read on his own and to react to books and mass media intelligently to understand and appreciate other people and other cultures.

In the acquisition of English as an effective tool for learning, a lot of complex skills are required. The Filipino child has to develop skills in listening, speaking, reading and writing. These are difficult enough to develop in a first language; they are real challenges in a second language. Take listening, for instance. The child has to learn to distinguish the sounds in L_2 different from his own L_1; he has to listen with understanding to what he hears in the new language. Again, speaking requires other skills like producing the phonemes of L_2 correctly and speaking with the correct stress, juncture, and intonation. The other two skills, reading and writing also call for other complex sub-skills.

The singling out and evaluation of these sub-skills can be a complicated process for the uninitiated teacher. Which listening skills, for instance, are critical in the learning of L_2 and how are they to be evaluated? What in oral language must be evaluated and how does one go about it?

Most English teachers in the Philippines, like perhaps other teachers in Southeast Asia, are content-oriented or syllabus-conscious. To them pencil-and-paper tests are the best instruments for evaluating coverage of a syllabus. They will readily translate into an objective test all items in the syllabus that can be measured in this manner. If phonology is not included in their test, it is because they have no idea of how to test auditory discrimination, aural comprehension, or oral production. Even if they recognize the need to evaluate listening and speaking, they would guess somehow that tests of these skills would be difficult to prepare and administer. The first big problem in the evaluation of English teaching is the lack of know-how among English teachers. They have to be trained to identify what features should be tested in language and how to test them.

It would be easier for teachers to measure the results of English teaching if there were available linguistically-prepared, reliable and valid tests after which they could pattern their own teacher-made tests. Better still, if such tests were available on the national level, they could be used for the improvement of English instruction all over the country. We have to continue to work harder to improve our achievement tests in reading and language to make them valid instruments for measuring over-all achievement in English. This means testing performance in English in all its aspects: pronunciation, aural comprehension, oral production and expression, reading comprehension, accuracy in writing, and such composition skills as unity and coherence. The latest national achievement tests in language and reading, series 1965, have been prepared in the "traditional" manner – "traditional" as used here meaning strictly pencil-and-paper testing or evaluating only those features of language that can be measured through a written test. The language test for Grade IV covers correct usage (grammar and vocabulary) and the writing of sentences as responses to questions and as context for the use of certain words. The language test for Grade VI includes some dictation and spelling in addition to correct usage. The reading tests which are patterned after reading tests in the United States, include word meaning, sentence meaning, and paragraph meaning.

Some pioneering work in modern language testing is being done in the Elementary English Section of the Bureau of Public Schools. They have just recently validated an achievement test in Grade IV language which has the earmarks of linguistic sophistication.

Compared with the Bureau national achievement tests, the periodic tests prepared on the division level by English supervisors or teachers who have had training in TESL show more linguistic orientation. The high school tests for example, evaluate oral production through partial production tests on segmentals, stress, and intonation. Some of the tests on structure are situational in that they require responses to situational questions or sequence sentences in short dialogues. The tests also show the test constructor's knowledge of contrastive analysis as shown in his choice of questions on structure which include critical items in word order, agreement, and tense. And last, but not least, the tests attempt an evaluation of simple writing skills.

Some local training institutions — the University of the Philippines, the Philippine Normal College, and the Ateneo de Manila University, all of them having TESL graduate training programs — have developed very good language tests which, however, have had only limited use. The Philippine Normal College leads with a total of around 21 tests submitted as graduate dissertations by teachers on scholarship grants. These tests have been validated and used in the schools or divisions where these teacher grantees are teaching.

The dissertations range from tests on specific areas of structure – e.g. prepositions – to tests on general achievement in language. There are also several types – diagnostic, achievement, and proficiency – and tests for various grade levels in the elementary and high school. There are tests in language and tests in reading. The University of the Philippines has tests of two types: diagnostic and achievement.

The best scientifically prepared tests in the Philippines today are the achievement tests used in the Rizal Experiment prepared in 1959 – 1964 by the Research Division of the Bureau of Public Schools and the Philippine Centre for Language Study under the guidance of Dr. Frederick Davis of the University of Pennsylvania. The series consists of language aptitude tests, an English proficiency test, and achievement tests in language and reading. The English proficiency test is especially note-worthy.

At this juncture it would seem that there is no lack of reliable and valid language and reading tests in English in the Philippines. These linguistically prepared tests, however, are in the form of unpublished theses and cannot be reproduced without the permission of the university or college where they were prepared. It would take, however, more than one of these tests to constitute a reliable measuring instrument for the evaluation purposes of the Bureau of Public Schools. We need to make our own tests which can be used on a national level for the improvement of English instruction. To make good tests we need technical know-how. We need time, released time, so that our test constructors can concentrate on test preparation without being bothered by other duties. We need the assistance and expertise of the SEAMEO RELC — its research facilities, and above all its bold pioneering spirit to keep us going forward. All this could be done with trained yet humble mind; as the great 16th century poet William Cooper once said "Knowledge is proud that he has learned so much; wisdom is humble that he knows no more".

Distinguished delegates to this scholarly gathering, this in brief is the status of English teaching and testing in the Philippines. We realize our limited resources and the priorities we have to undertake in the area of language teaching. We are happy that in this market place of ideas you have invited us to participate. We shall continue to keep pace with the progress of our neighbouring countries in Southeast Asia so that hand in hand we can move onward in the pursuit of knowledge and wisdom so that there will be better communication among nations and finally understand each other as brothers in a common search for peace and happiness among nations on this face of the earth.

ENGLISH LANGUAGE TESTING IN SINGAPORE

by NG KUEN SEONG

1. Background

Singapore is in a unique position where the teaching of English is concerned. The entire school-going population learns English, either as a medium of instruction or as a second language.

In the first instance, the teaching of English takes place in English-medium schools where total exposure time is approximately 30 hours per week, including participation in extra-curricular activities. The majority of the learners are learning English as a second language in so far as it is not their mother tongue.

In the latter case, English is literally taught as another language, a subject in an environment where most of the rest of instruction is undertaken in Malay, Chinese or Tamil. In terms of exposure to, and the use of the language, it is the teaching of English as a foreign language, for the learners hardly use or hear English outside the 6 x 30- or 40-minute periods per week. Although in the non-English primary schools, Mathematics and Science are now taught in English for another 8 periods a week (thereby more than doubling exposure time), it is still too early to assess the effects as the programme has just been implemented on a national scale for five months.

Officially, however, the term — the teaching of English as a second language — applies to the teaching of English in Malay, Chinese and Tamil streams or schools in Singapore.

2. English Language Examinations

In both our English and non-English medium schools, English is taught from the first year, for 6 years in the primary school and another 4 to 6 years in the secondary school. Periodic tests are held according to the programmes of individual schools to determine progress and promotion to the next class. There are also two examinations conducted in the middle and at the end of the year respectively. State examinations are administered at the end of the 6th, 10th and 12th year. These examinations are known as the Primary School Leaving Examination, the Cambridge School Certificate or Secondary IV Examination, and the Higher School Certificate Examination respectively.

There are other examinations at college or university level as well.

3. The Primary School Leaving English Examination

This is a local examination conducted by the Ministry of Education in Singapore for Primary VI pupils in all four language streams — English, Malay, Chinese and Tamil — of Government and Government-aided schools. Last year, approximately 75,000 candidates entered for the examination, an increase of 5.5 per cent over the 1968 figures. Of these, 64.72 per cent were from the English stream, the rest from the non-English streams. 53.4 per cent of the candidates were given places in secondary schools.

All the candidates sit for five compulsory papers: First Language (medium of instruction), Mathematics, History and Geography, Science and a Second Language (English, Malay, Chinese or Tamil by choice). Promotion to secondary schools is based on overall performance but the First Language paper carries a weighting of two units while the others carry only one unit each.

A table of specifications is issued to primary schools at the beginning of each school year to inform principals and teachers of the objectives of the examination and the break-down of the syllabus in terms of topics.

According to the table of specifications for English in English-medium schools, "the examination in English will be based on the whole of the syllabus for the teaching of English in primary schools issued by the Ministry of Education. Every section of this paper will attempt to test the candidate's knowledge of current English usage as applicable to present-day Singapore, and his ability to use the language effectively at a level that can be reasonably expected of an eleven- or twelve-year-old. A knowledge of grammatical terms will not be required of the candidate."

The same conditions obtain in our English-as-second-language environment except that the examination is based on the last two years of the primary syllabus for English in Malay, Chinese and Tamil schools

A panel of five or six language specialists and practising teachers with many years of teaching experience and training in examination techniques share the responsibility for the construction of each test item. After many sittings during which there is free exchange of views and frank criticism of the quality of each test item, the papers are moderated by a qualified person and approved by the Chief Examinations Officer before the tests are administered.

Among other qualifications, setters are supposed to have:

1. knowledge of the syllabus,
2. knowledge of the breakdown of the syllabus into parts and the detailed objectives which are to be tested in the paper as a whole and in the different questions,
3 judgement of what levels the average candidate should be capable of attaining, and
4. acquaintance with previous question papers, performance of candidates in previous examinations, criticisms of the chief examiners of previous question papers, etc.

Ideally, there should be a pre-test on a sample group of pupils but, owing to administrative problems, it has to be put aside. However, from the item analysis of the performance of each examination, it is possible to distinguish between the quality of one item and that of another, and this provides guidance in the setting of the following year's paper. Other factors considered are relevance, balance, objectivity, difficulty, and most important of all, reliability and validity of the tests.

4. Administration

English-stream candidates sit for two English papers on the first day of the 3-day period of the examination during which all primary schools are closed. There is a break of only 5 minutes between Paper I — Composition — and Paper

II — Language. Since the machinery for the Primary School Leaving Examination has been well established (it started in 1960), there are no administrative problems except the standardisation of marking of the Composition paper.

Non-English stream candidates sit for only one paper which tests composition, comprehension, vocabulary and usage. Since all the questions are multiple-choice types, no extra standardisation of marking apart from the marking scheme is needed.

5. General description of the English papers for the P.S.L.E.

For the English-medium candidates, two papers are set: Paper I which tests Composition and Paper II, Language. The Composition test carries a total of 50 marks out of a total of 200. The question for this paper provides guidelines for the candidates to follow. In the Language paper, the majority of the questions are multiple-choice ones. Matching and open-ended questions are also included. In testing proficiency in English, both the essay-type and the objective-type questions are used to combine the advantages of both. The essay or traditional type is a valuable supplement to the objective type as it calls for fluency of expression, individuality and originality. It further ensures that adequate opportunities are given to children in primary schools to practise writing in continuous prose. Since the inception of the Primary School Leaving Examination in 1960, objective testing in English has been used to cope with the increasing numbers of examinees and to ensure efficient and reliable scoring.

For Paper I, the candidate is expected to write a composition of at least 150 words within 40 minutes. The subject set tests the candidate's ability to describe people and places, narrate events and happenings and, to some extent, express personal feelings and opinions. In order to improve the reliability of essay test scores, questions are specific and accompanied by directions or guidance for the candidate. In Paper II, there are 140 items to test vocabulary, usage and comprehension in 1 hour 55 minutes. The most widely used form of objective testing is the multiple-choice, which calls for the most suitable or correct answer out of a list of four alternatives. Experience and item analysis have shown that the number of items is adequate as almost all candidates have had time to attempt all of them. There is no oral test.

The English-as-a-second-language paper tests 110 items in 75 minutes. 70 of the items test composition and comprehension with picture cues. All the items are given with a choice of 4 alternatives.

More important than the written paper is the oral test for non-English-medium pupils, which is administered in June-July each year for 3–5 days. At present, the test consists of 2 parts — reading of a passage of approximately 80 words, and answering 5 questions based on a picture.

6. Marking

As for the examination, the primary schools are closed for 3 days for the manual scoring of all the scripts. Markers are recommended by primary school principals and they fall into two categories for the English stream: General and Composition markers. The latter are teachers of English. Markers for the Composition paper are thoroughly briefed: scoring is guided by explicit detailed instructions on the allocation of marks for different aspects of the answer and the deduction of marks for different kinds of errors; and, since marking is centralised, scripts are constantly being moderated by the moderators on the spot. Any deviation is pointed out by the supervisor and adjustments are made immediately. Objective-type questions are marked and checked by markers, working in pairs, under proper supervision. Marks are finally processed by a computer.

7. **Item Analysis**

A small-scale item analysis of the English Language paper has provided some indication of the suitability of the items set. It is hoped that, with the co-operation of the Research Division, a large-scale analysis of all subjects will be carried out this year.

8. **The Cambridge School Certificate English Examination**

The Cambridge School Certificate Examination marks the end of the fourth year of secondary education in the English stream. The English papers, also two in number, are set and marked by the Cambridge Examination Syndicate.

The English Language paper is a compulsory paper in the School Certificate examination in the English stream. Its object is to test the candidates' ability to write in the English Language correctly; the main emphasis is on linguistic achievement.

English Language (Syllabus X) consists of 2 papers of 1½ hours each, with equal maximum marks allotted to each paper. Candidates must take both papers.

In Paper I there are alternative versions, A and B, printed together as one paper. Alternative A contains a choice of alternative subjects for continuous composition. Material for some of these is provided in the question paper while others are general subjects. Alternative B, is intended to provide an alternative English test consisting of two short compositions rather than the single longer composition required in Alternative A, and to give candidates an opportunity to show their ability to expound factual matters with relevance, clarity, economy and accuracy. The paper is divided into 2 sections, both of which have to be attempted and which carry equal marks. Section (i) provides a choice of alternative subjects for composition, one to be attempted; section (ii) is a test of continuous writing, a great part of the material for which would be supplied in the question paper.

In assessing a Paper I script, examiners first place it in one of five main classes, A – E, according to its general linguistic merit. Then, with special reference to vocabulary, idiom, sentence structure, paragraphing, and links, the examiner places the script more definitely in a mark category (B^+, B, B^-, C^+, etc.) within the main class. Finally, a slight adjustment up or down is made for content and arrangement considered together; and a numerical mark is allocated.

Paper 2 consists of a passage or passages of prose upon which questions are set to test ability to understand the content and argument of a given text and to infer information and meaning from it. Ability to summarise may also be tested. In connection with this paper, a book is recommended for background reading on a topic similar to that which would provide the theme for the passage(s) on which questions would be set. However, the passage(s) would not be taken from the book recommended and the paper can be attempted without a knowledge of the book.

The Oral English test is compulsory for all candidates, and comprises 20 per cent of the overall marks for the English Language Paper.

The examination is in 3 parts with marks allotted as follows:

Reading 100		
	Pronunciation	50
	Fluency and Rhythm	50
Comprehension 50		
	General Comprehension	25
	Verbal Accuracy and Fluency	25
Conversation 50		

The Chief Oral English Examiner holds meetings with local Examiners before the examinations begin. The aim of these meetings is to obtain agreement upon a common standard of marking. Passages for reading and comprehension are set by Cambridge and attempts at standardising evaluation include samples on tape which are played to local examiners.

Literature in English

This is one of the 4 papers that can be offered on a group of subjects called General Subjects, the others being Religious Knowledge, History and Geography. This paper, formerly known as English Literature, was re-named Literature in English in 1968 to include the works of writers of Commonwealth and other countries.

Alternative syllabuses (i) and (ii) may be attempted; the candidate, however, may not combine a part of syllabus (i) with a part of syllabus (ii).

The First Alternative consists of one paper of 2½ hours. Candidates have to answer 5 questions and have to offer the prescribed Shakespeare play and 3 other texts including at least one of the prescribed plays or anthologies of poems, and one of the prescribed novels.

9. **The Secondary IV Examination** (Malay, Chinese or Tamil)

The English Language papers for this level are set and marked by the Ministry of Education in Singapore. For the three streams, there is a common 2½-hour paper with 4 sections: Composition (about 150 words), Letter-writing, Comprehension, and Structure and Usage. The first two sections are essay tests while the latter 2 sections comprise traditional tests on comprehension and grammar (or structure) and usage, and multiple-choice type questions.

For the Malay-stream candidates, there is a Lower English paper inherited from the Malaysia Certificate of Examination set-up. This is a pass-level paper which used to be set and marked by the Cambridge authorities. From last year onwards, it became part of the local examination system.

The test consists of one 2¼-hour paper which is largely an objective type. Section A tests continuous prose writing with the help of a plan. The other sections are similar to those for the Secondary IV examination except that all the questions are multiple-choice ones.

Markers come from the ranks of practising teachers of English as a second language with experience in marking such scripts. A comprehensive marking guide provides detailed instructions for allocating or deducting marks. All markers and moderators attend a compulsory briefing session before marking commences.

There is no oral test at this level.

10. The Cambridge Higher School Certificate English Examination — General Paper and English papers

The General Paper

All candidates for the full HSC are required to reach a satisfactory standard in the General Paper, a pass in which is recorded as a pass in a subsidiary subject. However, an HSC certificate may be awarded to a candidate who fails by a small margin the General Paper but who satisfies other general conditions.

One 2½-hour paper is set to test the understanding and use of English and the extent to which the candidate has achieved a maturity of thought appropriate to Sixth-Form Students in their second year. This paper is not primarily a test of general knowledge.

The paper is divided into 2 sections, A and B. Section A consists of topics for composition in 3 sub-sections which include historical, social, economic, political and philosophical topics as well as topics on science, mathematics, geography, literature and language, and art and crafts. Answers to questions in Section A would normally be 500–800 words in length.

Section B tests comprehension of an English prose passage as a whole and in detail; the ability to re-express in continuous form material supplied in the paper; the knowledge and understanding of common English usage. Three questions are set, based on information given in the question paper; one tests comprehension, the other two vary and may be based on information given in the form of statistics or diagrams or may take the form of tests of logical and scientific reasoning.

Candidates in Singapore are expected to answer one question from each of Sections A and B.

English (Principal level)

To obtain a pass at principal level in English in the Cambridge Higher School Certificate examination, candidates are required to offer three of eight papers; including at least one of Papers 2 (i.e. Shakespeare) and 3 (i.e. Chaucer and other Major Authors). Papers are of 2½-hour duration.

Paper 1 (Composition and Comment) includes passages for comment and appreciation. The paper is divided into 2 sections; section (*a*) consisting of a passage of modern prose for summary or precis, and either a passage of 16th- or 17th-century prose for translation into modern English, or a passage of modern English for detailed comprehension.

Section (*b*) consists of a passage or short passages of literary prose or verse for exposition, comment or comparison.

Paper 2:	Shakespeare — candidates must offer a minimum of two plays of which one must be a tragedy and the other either a comedy or a history.
Paper 3:	Chaucer and other Major Authors.
Paper 4–8:	are set on the following periods, and candidates must offer a minimum of 4 books.
Paper 4:	The period 1550–1660.

Paper 5:	The period 1660–1780.
Paper 6:	The period 1780–1832.
Paper 7:	Literature of the Victorian Age.
Paper 8:	English Literature since 1900.

Any one of the above 8 papers may be offered as a subsidiary subject. There is also a special paper in English at subsidiary level.

The Syndicate issues, at intervals, reports based on the work of candidates in the various papers. These reports are prepared for schools by the Syndicate's Examiners and are meant to be of practical help to teachers and students.

11. **The Higher School Certificate Examination** (Malay or Chinese)

The Malay-stream candidates sit for a General Paper in English. It is also set and marked by Cambridge.

The Chinese-stream candidates are offered two optional papers — Paper I, which is like the Cambridge School Certificate English Paper; and Paper II, a Literature paper like that of the Cambridge HSC paper. Paper I is compulsory for candidates who wish to further their studies at the University of Singapore, Nanyang University, or the Polytechnic; and Paper II for those who wish to read English at the University of Singapore. Both are very traditional.

An oral test is administered at this level. Although it is set by the Ministry of Education, it is very much like that of the Cambridge School Certificate oral test which is far too subjective and tests hardly anything.

12. **The Future**

It has been officially announced that from 1971 onwards, all papers set in the English language will come from Cambridge under the new partnership, the Singapore-Cambridge General Certificate of Education Examination. Other subjects to be examined in Chinese, Malay and Tamil will be the responsibility of the Singapore Ministry of Education. Only one certificate — the Singapore-Cambridge General Certificate of Education will be awarded to successful candidates.

THE OBJECTIVES OF ENGLISH LANGUAGE TEACHING AND THE STATUS OF ENGLISH TESTING IN THAILAND

by BANLUE TINPANGA

1. **Background Information**

The educational system of Thailand may, to serve the purpose of this presentation, be sketched as follows:

Elementary education	
Lower elementary	4 years
Upper elementary	3 years
Secondary education	
Lower secondary	3 years
Upper secondary	2 years
(2-year teacher training, lower certificate level)	

Higher education

- — non-education bachelor's degree programs
- — 2-year teacher training (higher certificate level)
- — 4-year bachelor of education program
- — 2-year bachelor of education program (sequel to the higher certificate level)
- — graduate programs

Within this educational system, English is the required foreign language that extends over the longest period of time — from the 5th grade through to the university. It is believed that English will facilitate Thai children's living, although it has never been ascertained statistically on a nationwide basis how much English is used in the ordinary life of the Thai people as a whole.

At the level of higher education, however, the need for English is quite self-evident. Thai college students cannot proceed very far with their acquisition of knowledge before they are compelled to turn to English books. But there is more or less a consensus in Thailand that secondary school graduates, after 8 years of English, are ill-equipped to handle the language at the level required for their college education. More will be elaborated on this point later.

2. **The Objectives of English Teaching in Thailand**

Objectives

The objectives of English teaching in Thailand may be reduced to the following:

1. To develop facility in the use of English; and
2. To develop an understanding of the cultures of the English-speaking peoples.

Obviously these two objectives involve numerous activities necessary for their realization. They are also set in various stages of difficulty and depth.

The first objective inevitably entails the practice of the four language skills: speaking, listening, reading and writing. This is done from the beginning of the English instruction at which stage the goal is for the pupils to be introduced to the language and to learn English in its simplest form, while at the level of higher education students are expected to understand and express not only simple things such as an observation on the weather but also more complex ideas. Throughout the English program a distinction is made in the content as to which is intended for recognition and which for production. The use of such difficult tenses as the past perfect, for example, is postponed until the upper secondary and even the university level, whereas its recognition comes earlier because the tense is sometimes found in textbooks for the lower levels.

The second objective obviously owes its presence to the belief that language is part of a culture, although it is not definitely known how well Thai students have learned English with or without the knowledge of the cultures of its peoples. (The term "culture" here is not meant to be identical with the ideal language learning situation where only the target language is used.) At any rate, the part on culture has a secondary, complementary role in the English syllabus in Thailand, as it perhaps should.

The objectives in relation to the English teaching and learning situation in Thailand

Fulfilment of the objectives

The teaching and learning of English in Thailand can be regarded as having improved from the situation of 15 years ago. At that time a great deal of the teaching and learning was *about* rather than *of* English. In comparison, the students of the present day are more confident of their ability to speak the language, and speaking is the skill that has improved the most, thanks to the new awareness that a language is better spoken than merely talked about endlessly. The teachers of English have, of course, improved in both knowledge and technique. The students themselves have also become less reticent than their traditionally reserved predecessors. Perhaps this trend toward freer expression on the part of the students enables them to learn a foreign language, especially the oral part, with more degree of success than before.

This improvement, however, is highly relative. As stated earlier, the success in English of our secondary school graduates is far from satisfactory — in relation to the objectives, that is. Complaints can frequently be heard from several quarters: from university instructors that secondary schools poorly prepare their students in English; from secondary school teachers that elementary schools have not laid a proper foundation for the students' English; and finally from elementary schools that teacher training institutions have not given prospective teachers adequate preparation in the teaching of English. Certainly the question is not who is to blame. The point here is that the situation seems to indicate a re-examination of the entire English program. A study will be needed to determine if the set objectives are too ideal for Thailand, where English has the status of a foreign language far removed in its use from the native tongue. It seems desirable that clear definition be made as to what level of attainment is intended for which group of people or students.

The question of economy in ordering an instructional program

Thailand, like her neighbours in Southeast Asia, can benefit by observing the rule of economy in its undertakings. As pertinent to the English language program, it must be realized that several hundreds of hours are spent on learning

English with the assumption that English has been learned. But experience has indicated that these hours are not always well spent. Here again a study of a kind seems to be in order so as to find out exactly what is an optimal number of hours to be devoted to English for a certain level of attainment, and where lies the most suitable starting point for English *in the Thai school system*.

3. **English Language Testing in Thailand**

Forms and purposes of tests

The present forms of English tests in Thailand are predominantly of the objective type: multiple choice, completion, blank filling, substitution, and so on. The essay type is used comparatively little, so the use of free composition in a test occupies only a small portion. It is used mostly at the university level, with the result that the students' composition is usually chaotic. Aside from the mechanics of language, it is suspected that the students' poor ability to express ideas is carried over from Thai, where they do not fare much better.

The above forms of tests are used for various purposes: to test the skills in reading, writing, comprehension, and the knowledge of the structure of the language. The oral skill is rarely tested formally. It is left to the teachers to measure in class the students' oral progress.

Types of testing

Broad-scope testing

This type includes centrally made examinations such as the nationwide examination at the end of secondary school, the university entrance examination, and the regional examination at the end of the 3rd year of secondary school. English is one of the subjects examined at these examinations. The nationwide examination at the end of the 5th year of secondary school consists of 3 papers on English, of which the first paper is taken by all secondary school leavers. The second paper tests the knowledge of the readings, in which arts and science students choose to answer questions on different books. To do this paper requires some memorization, no doubt, and some understanding of stories from the English-speaking world. The third paper is for arts students only. It calls for deeper knowledge of language structure than does the first paper. The university entrance examination also follows the same line in regard to the requirements for arts students and for non-arts students.

The regional examination consists of contents in the same categories as the nationwide examination, but the papers are usually divided into comprehension, expression, reading and grammar. It is not as fixed in form as the nationwide examination, since each of the 12 educational regions in the country makes its own examination.

Limited-scope testing

Included in this category are various entrance examinations excluding the university entrance examination, and regular tests in courses of study. Entrance examinations at the College of Education and teacher training schools are prepared by the institutions themselves. Students entering the 1st and 4th years of secondary school as well as those entering teacher training schools are required to take an entrance examination in English as a subject. The entrance examination at each higher level aims to test the knowledge of English that has been supposedly acquired throughout the years. Again it is not possible to test the oral skill here.

These entrance examinations are national in the sense that they make up a uniform practice, although their forms and standards may vary.

Tests in each course of study are perhaps self-explanatory. They are given both formally and informally, and it is here that the oral skill is mostly tested. Each course being a part of an integrated whole program, it is hoped that a cumulative grade point derived from all the courses in the program will reflect the general ability in English of each student. This type of testing is most frequent in teachers colleges and other higher educational institutions, although informal testing of the oral skill is done through all the levels.

4. **Problems and Solutions**

Problems

Lack of balance of the language skills in testing

Because of the lack of speaking proficiency on the part of many Thai teachers, the oral skill is apt to be neglected even in a class situation where informal oral testing can be carried out. Further, the other skills also suffer because of this shortcoming.

Lack of knowledge of testing technique

Most tests and examinations involve the placement and the end-result of students' English ability, but perhaps there is not yet enough testing of students' progress while an English course is moving along. This progress testing could be of great help in the improvement of each individual student.

Lack of concerted efforts and an information clearing and distributing centre

This topic hardly needs elaboration. There are no doubt separate efforts being made in the field of teaching English in Thailand and, if well coordinated, they can be turned into a very dynamic and beneficial force. With effective coordination, the English programs at all the levels will also be blessed with a clear and firm sense of purpose.

One solution to the first two problems has been the offering of in-service training programs to remedy specific problems of teachers. An in-service training program has proved to be most effective when it is an intensive one given over a reasonable length of time. Also, teachers are allowed leaves of absence to continue their studies.

As for the third problem, a solution so far has been the recently established English Language Centre of the University Development Commission. It remains to be seen how this problem can be tackled further.

ENGLISH LANGUAGE TESTING AT HIGH SCHOOL LEVEL IN VIETNAM

by DUONG THANH BINH

Every year, around November, the Director General of Primary and Secondary Education in Vietnam orders the principals of all public secondary schools to inform their teachers of the preparation of tests for the two national examinations, the so-called Baccalaureate Part I and Baccalaureate Part II. These two examinations are held twice a year and are provided for those students who have completed six or seven years of high school studies.

The above order requests those teachers who teach classes that are going to sit for the examination to propose two tests and send them to the Testing Department of the Ministry of Education. The teachers are also instructed to propose tests which are general enough to cover the whole program, appropriate to the level of the students, and to the time limit of the test. Teachers of English receive specific instructions concerning the number and types of questions given, and the number of marks given to each question. From the proposed tests, the Testing Department will select those which are most appropriate, and use them as the actual tests for the year's examination.

The objectives of the national examination are to measure and check the learning and comprehension of the students about the subjects they have studied in high school. With English, for instance, the students are tested on their skills in pronunciation, grammar, vocabulary, reading, and writing.

According to the English syllabus of the Ministry of Education, the objectives of teaching English to high school students are to enable them to understand, speak, read, and write good and fluent English, to understand and appreciate the customs and culture of the people whose mother tongue is English. To achieve these purposes, the *English for Today* series of the McGraw-Hill Company have been used as textbooks, and the Oral Approach with pattern drill techniques and audio-visual aids has been suggested to teachers of English as a good teaching method.

As in many other countries in the world, foreign language teaching and learning start at the high school level in Vietnam. The two foreign languages taught are French and English. Depending on the choice of the student upon entering high school, one of these languages will become his first foreign language, and the other will be his second foreign language. The length of time for the study of the first foreign language, either French or English, is 7 years, i.e., throughout the high school period. The study of the second foreign language takes place during the last three years of high school.

The Vietnamese high school is divided into two levels: the 4-year junior high school, and the 3-year senior high school. In senior high school, the studies are divided into four different sections: the Natural Sciences section or section A, the Mathematics section or section B, the Modern Languages section or section C, and the Classical Language section or section D. If English is the second foreign language, it is taught 3 or 4 hours a week in sections A and B, and 6 hours a week in section C.

As mentioned earlier, the Baccalaureate Part I examination occurs at the end of the sixth year of high school. The students have to pass this examination before they are allowed to continue their last year in high school and to sit for the Baccalaureate Part II examination. In Appendix *D* (p. 62) you will find some samples of the English tests that were used in the examination for the Baccalaureates Part I and Part II. You will notice the differences in the length of time given to complete the tests, and the differences in the content of the tests. The longer and more difficult tests are for those students who choose English as their first foreign language, and the shorter and easier ones for those who choose English as their second foreign language. If you look closely at the tests, you will also notice that the questions are divided into different types. There are questions on pronunciation, vocabulary, grammar, word formation, and translation. In the test for Sections A and B with English as the first foreign language, the emphasis is mostly on translation. In the tests for Sections C and D with English as the first foreign language, the emphasis is on composition or writing skills. In the tests with English as the second foreign language, there is no question on translation, and the marks are evenly distributed among all the questions.

If we take into consideration the fact that the students who take the tests have had 2 or 3 years of English as the second foreign language, and 6 or 7 years of English as the first foreign language, we can find that the content of the tests given is very limited. Although every aspect of language learning is being tested, the questions given are too simple and short to be able to measure truthfully the achievement of the students in their learning of English. And there has been no way to evaluate the validity of the tests. A few years ago, there were oral tests in addition to the written test, but due to the rapid increase in the number of students and the complication in the administration of such a test, this practice had to be discontinued.

Besides the differences in the length of time of studies, in the content of the tests, and in the time limit to complete the tests, the differences between English as the first and second foreign language are most noticeable in the percentage of marks given to the tests. In Vietnam, the marks are based on the decimal system. Every subject is graded with a maximum grade of 20 marks and a minimum of zero (0) mark. The average or passing mark is 10. If one test is considered more important than the others, it is given a larger share in the total marks. This is done by considering it equal to 2, 3, or 4 tests. In other words, it is given a coefficient of 2, 3, or 4. (Cf. Appendix *A*, p. 57). There are actually seven tests for each examination, but the total coefficient or marks are the same as if there were 14, 15 or 16 tests. For example, in order to pass the Baccalaureate Part II, sections A and B, the students need to have a minimum of 160 marks. Another characteristic of this system is that the marks of a student's test can be transferred to other tests if that test receives more than the necessary average marks. On the other hand, if one of his tests receives a zero mark, the student automatically fails the examination, even though he gets enough marks from the other tests to pass the examination.

In Appendix *B* (pp. 38–59) you will find statistics which show how many boys and girls sat for the 1969 national examinations, and how many of them passed. But these figures do not show how many of these boys and girls took English as their first or second foreign language, how many of them received enough marks for their English tests, what part of the test was the most troublesome to them, or whether or not the test was valid in measuring the achievement of the students. In fact, no studies have been made to judge the validity of the test from the point of view of construction, application, and evaluation.

Recognizing these shortcomings, the Ministry of Education has recently ordered studies to be made to help improve the curriculum as well as the methods of teaching, and the procedures and techniques of testing in the national high school examination.

Appendix A

COEFFICIENT OF THE TESTS IN BACCALAUREATE PART I AND PART II EXAMINATIONS

A. *Baccalaureate Part I*

	A	B	C	D
First foreign language	2	2	3	3
Second foreign language (Classical language for D) ..	1	1	2	3
Mathematics	2	4	1	1
Chemistry–Physics	3	3	1	1
Natural Sciences	3	1	1	1
Vietnamese	2	2	4	4
Civics, History–Geography	1	1	2	2
Total Coefficient ..	14	14	14	15

B. *Baccalaureate Part II*

	A	B	C	D
First foreign language	2	2	3	3
Second foreign language (Classical language for D) ..	1	1	2	3
Mathematics	2	5	1	1
Chemistry–Physics	4	4	1	1
Natural Sciences	4	1	1	1
Philosophy	2	2	4	4
Civics, History–Geography	1	1	2	2
Total coefficient ..	16	16	14	15

A = Natural Sciences Section
B = Mathematics Section
C = Modern Languages Section
D = Classical Language Section

STATISTICS OF THE RESULTS OF 1969 BACCALAUREATE PART I EXAMINATION

Examination Committee	Section	CANDIDATES TAKING THE EXAM			CANDIDATES PASSING THE EXAM			Percentage
		Boys	Girls	Total	Boys	Girls	Total	
Southern VN								
Independent ..	A	5,072	2,137	7,209	250	98	348	4.82%
	B	6,046	1,055	7,101	476	45	521	7.29%
	C	2,415	925	3,340	418	190	603	18.05%
	D	33	01	34	06	00	06	17.65%
Semi-Public and Private School ..	A	2,100	5,215	7,315	362	614	976	13.34%
	B	6,364	2,034	8,398	2,229	375	2,604	31.09%
	C	593	749	1,342	266	225	491	37.40%
	D	295	25	320	82	03	85	26.59%
Public School ..	A	1,449	3,840	5,289	645	1,533	2,178	41.19%
	B	3,105	902	4,007	1,823	398	2,221	52.88%
	C	74	150	224	37	77	110	49.56%
	D	01	00	01	00	00	00	00%
Total ..		27,547	17,033	44,580	6,689	3,548	10,237	22.96%

STATISTICS OF THE RESULTS OF 1969 BACCALAUREATE PART I EXAMINATION

Examination Committee	Section	CANDIDATES TAKING THE EXAM			CANDIDATES PASSING THE EXAM			Percentage
		Boys	Girls	Total	Boys	Girls	Total	
Central VN and the Highland								
Independent ..	A	1,935	743	2,678	171	82	243	9.07%
	B	3,714	375	4,089	575	53	628	15.35%
	C	1,044	290	1,334	365	72	447	32.75%
	D	90	00	90	22	00	22	24.44%
Semi-Public and Private School ..	A	1,025	1,021	2,046	134	118	252	17.20%
	B	2,401	353	2,754	479	49	528	19.17%
	C	205	128	333	103	55	158	47.44%
	D	15	00	15	14	00	14	93.33%
Public School ..	A	938	1,038	1,970	204	228	432	21.88%
	B	1,806	478	2,284	646	135	781	34.19%
	C	163	274	437	90	158	248	56.75%
	D	00	00	00	00	00	00	00%
Total ..		13,336	4,700	18,036	2,803	940	3,743	20.79%
Total ..		40,883	21,733	62,616	9,492	4,488	13,980	21.88%

Appendix C

STATISTICS OF THE RESULTS OF 1969 BACCALAUREATE PART II EXAMINATION

Examination Committee	Section	CANDIDATES TAKING THE EXAM			CANDIDATES PASSING THE EXAM			Percentage
		Boys	Girls	Total	Boys	Girls	Total	
Southern VN								
Public School ..	A	1,774	2,987	4,761	880	1,425	2,205	46.31%
	B	1,977	1,010	2,987	1,408	418	1,826	61.12%
	C	422	517	939	256	326	582	61.98%
	D	00	00	00	00	00	00	00%
Semi-Public and Private School ..	A	1,239	1,335	2,574	616	616	1,232	47.86%
	B	2,319	263	2,582	617	137	754	29.58%
	C	60	130	190	30	62	92	48.44%
	D	111	00	111	89	00	89	80.18%
Independent ..	A	1,038	929	1,967	194	219	413	20.99%
	B	2,270	130	2,400	420	76	496	20.66%
	C	895	410	1,305	188	279	467	35.78%
	D	65	00	65	47	00	47	72.30%
Total ..		12,170	7,711	19,881	4,745	3,458	8,203	41.36%

STATISTICS OF THE RESULTS OF 1969 BACCALAUREATE PART II EXAMINATION

Examination Committee	Section	CANDIDATES TAKING THE EXAM			CANDIDATES PASSING THE EXAM			Percentage
		Boys	Girls	Total	Boys	Girls	Total	
Central VN and the Highland								
Public School ..	A	570	681	1,251	225	312	537	72.30%
	B	1,304	889	2,193	492	354	851	38.80%
	C	401	417	818	233	220	453	55.37%
	D	01	00	01	01	00	01	100.00%
Semi-Public and Private school ..	A	250	470	720	56	101	157	21.80%
	B	245	170	415	73	49	122	29.39%
	C	80	70	150	34	38	72	48.00%
	D	19	00	19	16	00	16	83.15%
Independent ..	A	757	370	1,127	87	66	153	13.57%
	B	493	177	670	91	40	131	19.55%
	C	173	98	271	50	40	90	33.21%
	D	48	06	54	18	00	18	33.33%
Total ..		4,341	3,348	7,689	1,376	1,225	2,601	33.82%
Total ..		16,511	11,059	27,570	6,121	4,683	10,804	39.18%

Appendix D

Sample Tests

1968 Baccalaureate Part I Examination.
English Test.
Section C (the first foreign language).
Section D (the only living language)
Time: 3 hours. *Coefficient:* 3.

Read the following passage carefully and then answer the questions:

THE LOVE OF THE SEA

He — man or people — who putting his trust in the friendship of the sea, neglects the strength and cunning of his right hand, is a fool!

As if it were too great, too mighty for common virtues, the ocean has no compassion, no faith, no law, no memory. All the tempestuous passions of mankind's young days, the love of loot and the love of glory, the love of adventure and the love of danger, with the great love of the unknown, and vast dreams of dominion and power, have passed like images reflected from a mirror, leaving no record upon the mysterious record of the sea. Impenetrable and heartless, the sea has given nothing of itself to the suitors of their precarious favours. Unlike the earth, it cannot be subjugated at any cost of patience and toil. For all its fascination, that has lured so many to a violent death, its immensity has never been loved as the mountains, the plains, the desert itself have been loved.

J. CONRAD

Questions:

1. Put the accent mark *over* the main stressed syllables of the following words: TO NEGLECT, MEMORY, TEMPESTUOUS, ADVENTURE. (1 mark).

2. Use appropriate tenses in the following sentences:
   (*a*) The writer (die) years before his books (become) famous.
   (*b*) He writes as if the sea (be) a human being. (1.5 marks).

3. Write sentences of your own with: PRECARIOUS, FASCINATION. (2 marks).

4. (*a*) So far as Conrad was concerned, why shouldn't we put our trust in the ocean's friendship? (1.5 marks).
   (*b*) What are the common virtues according to him? (1 mark).
   (*c*) Did human passions mean anything to the sea? (1 mark).

5. Write a composition on *one* of the following topics: (12 marks).
   (*a*) The "fascinating" attraction of the sea.
   (*b*) You spent your last summer vacation at the sea-side. Write a letter to a friend of yours telling him (or her) how you spent your time, what you and other people did, etc. (Don't sign your name at the end of the letter.)

1969 Baccalaureate Part II Examination.
English Test (first foreign language).
Section A and B.
Time: 2 hours. *Coefficient:* 2.

Read the following passage carefully and then answer the questions:

If all mankind made it possible for Borman, Lovell and Anders to reach the moon, men the world over genuinely shared in their adventure. It was the farthest, fastest and most daring journey ever made — "One of the great pioneering efforts of mankind," said Dr. Thomas O. Paine, acting administrator of NASA. And of all the accomplishments, four stood out:

First, never before had man seen what Borman, Lovell and Anders saw as they circled 70 miles above the moon. Their feat gave the world a new and humbling perspective: looking back at the clouded earth from the moon, it was impossible to discern if life exists on this planet.

Second, their close-in inspection revealed a lonely, bleak and oppressive moonscape. "The color of the moon looks like a very whitish gray," reported Anders . . .

Third, the flight was a pathfinder for the Apollo mission this summer that is intended to land Americans on the moon.

Fourth, Apollo 8 was a triumph of U.S. technology. The 363-foot-tall rocket and ship, with no fewer than 3.5 million working parts, performed flawlessly . . . Launching was only six-tenths of a second late; lunar orbit was just one-half mile off; splash-down came within 7,000 yards of the carrier . . .

From *Newsweek*, January, 6, 1969.

Questions:

1. Put an accent mark *over* the stressed syllable of: mankind, genuinely, adventure, administrator, accomplishment, perspective, impossible, technology. (2 marks).

2. Give the definition of: ACCOMPLISHMENT, MOONSCAPE, WHITISH. (3 marks).

3. Change to indirect speech:
"The colour of the moon looks like a very whitish gray," reported Anders. (1 mark).

4. Give the above passage an appropriate title. (2 marks).

5. Why did the three astronauts' feat give the world a new and humbling perspective? (2 marks).

6. For what reasons is Apollo 8 considered as a triumph of U.S. technology? (2 marks).

7 Translate into Vietnamese: "First, never before . . . Americans on the moon." (8 marks).

1969 Baccalaureate Part II Examination.
English Test.
Section C (the first foreign language).
Section D (the only living language)
Time: 3 hours. *Coefficient:* 3.

Read the following passage carefully and then answer the questions:

Through Chicago passes the trade of the Lakes and the trade of the great West. The more wheat comes out of the soil and the bigger the litter of pigs, the larger grows Chicago — the highest example in the present world of the tendency of modern men to cluster into towns. The site of it is low and flat. The shores of the lake on which it stands are low all round, and we shivered as we were looking at the docks in the nipping wind which blew across from Canada. The city is impressive from its vastness, as the American rivers are impressive; one street, I was told, was many miles long. The stores are gigantic; the shops, etc., are large, and as if struggling to be larger, from the amount of business going on in them. If a house is placed inconveniently, they lift it on rollers and move it bodily from one spot to another, while the occupants sleep and eat and go on with their employments as if nothing was happening. I myself saw a mansion travelling in this way without the help of an Aladdin's lamp. To strangers, especially British strangers, the attractive sight in Chicago is the pig-killing. Five thousand pigs in a day, I believe, are despatched, cut up, and made into hams and bacon ready for packing. For myself, I had no curiosity to see pigs killed, nor, indeed, much for Chicago itself, beyond what a walk would satisfy, for towns of this kind are like the articles in which they deal — one part is just like another, you examine a sample and you multiply this by the dimensions.

J. A. FROUDE, *Oceans*

Questions:

1. (*a*) Put an accent mark *over* the main stressed syllable of the following words:
Tendency, impressive, believe, curiosity. (1 mark).

(*b*) Write a noun from: impressive, attractive, to believe, to satisfy. (2 marks).

(*c*) Turn into exclamatory sentences (2 forms):
"The stores are gigantic." (1 mark).

(*d*) Give the contrary of: modern, across. (1 mark).

2. Translate into Vietnamese:
   "The city is impressive . . . Aladdin's lamp." (3 marks).
3. Choose one from the two following topics:
   (*a*) Which do you prefer, living in the country or in a big city like Chicago? State your reasons.
   (*b*) River-side pleasures. (12 marks).

1968 Baccalaureate Part I Examination.
English Test (second foreign language).
Sections A and B.
Time: 1 hour. *Coefficient:* 1.

Read the following passage carefully and then answer the questions:

A GOOD CITIZEN KEEPS HIS CITY CLEAN

A family that lived in New York went to the park on Sunday. There were several children in this family. They played on the grass in the park. They thought it was pleasant to eat their lunch there, also. After lunch, the father read while he rested under the trees. At the end of the day, they all went home. The ground was covered with papers which they had thrown about. Other persons did the same thing. On Monday morning the park looked very ugly. Paper and dirty things must be placed in the baskets which are in the parks and along the streets. A good citizen must help to keep his city clean.

Question:

1. Group into sound columns: p*a*per, th*ou*ght, *ea*t, pl*a*ce, Y*or*k, tr*ee*, d*ay*, m*or*ning, str*ee*t. (3 marks).
2. Turn into the Simple Present Tense:
   (*a*) The father read while he rested under the trees.
   (*b*) They thought it was pleasant to eat their lunch there. (4 marks).
3. Turn into the Passive Voice:
   Other persons did the same thing. (2 marks).
4. Give the contrary of: pleasant, after, clean, big. (4 marks).
5. Why did the park look ugly on Monday morning? (3 marks).
   What are the baskets in the parks used for? (4 marks).

1969 Baccalaureate Part I Examination.
English Test (second foreign language).
Section C
Time: 2 hours. *Coefficient:* 2.

Read the following passage carefully and then answer the questions:

A RAINY DAY AT THE SEA-SIDE

There were only two Americans stopping at the hotel. They did not know any of the people they passed on the stairs on their way to and from their room. Their room was on the second floor facing the sea. It also faced the public garden and the war monument. There were big palms and green benches in the public garden. In the good weather there was always an artist with his easel. Artists liked the bright colours of the hotels facing the gardens and the sea. Italians came from a long way off to look up at the war monument. It was made of bronze and glistened in the rain. It was raining. The rain dripped from the palm trees. Water stood in pools on the gravel paths.

HEMINGWAY, *In Our Time*

Questions:

1. From the passage pick out 4 words containing the sound /ei/ as in 'R*ai*n' and two words containing the sound /ae/ as in "passed". (3 marks).
2. Find in the passage the opposite of: PEACE; PRIVATE; CONTINUING; DISLIKED. (4 marks).

3. Re-write the sentence: "THEY DID NOT KNOW ANY OF THE PEOPLE THEY PASSED." replacing *any* by another word. (2 marks)
4. Explain the difference between "TO STOP TALKING" and "TO STOP TO TALK". (2 marks)
5. Where was the two Americans' room in the hotel and what did it face? (2 marks).
6. Who often came to the public garden in the good weather? (1 mark).
7. Do you like the rainy season or the dry season? Why?
   (Write at least 2 sentences.) (4 marks).
8. Give the three forms of the following verbs.
   TO KNOW; TO STAND; TO STOP; TO COME. (2 marks).

1967 Baccalaureate Part II Examination.
English Test (second foreign language).
Sections A and B.
Time: 1 hour. *Coefficient:* 1.

Read the following passage carefully and then answer the questions:

WHAT ARE STARS?

Do you like to count the number of stars in the sky on a clear dark night? You will soon grow very tired for you can see about two thousand stars with your naked eyes. There are millions and millions of other stars, but they are so far away that you would need a telescope to see them. Even the number of stars that may be seen through telescopes grows larger and larger from year to year. This is because scientists can see more stars as they make better telescopes.

We are told that the stars are always moving in space. A few of them travel in groups, but most of them move alone. The universe is so very great that though the stars travel on and on, they may never come near each other.

Questions:

1. Group into 4 sound-columns according to the pronunciation of the stressed vowel sound in each word:
   n*u*mber, th*ou*gh, *e*ven, gr*ea*t, al*o*ne, n*a*ked, *o*ther, *ea*ch. (4 marks).
2. Give: (*a*) An antonym for: more, better.
   (*b*) An adjective from: universe, space.
   (*c*) A verb from: large, dark. (3 marks).
3. Turn into the passive voice: "Scientists can see more stars as they make better telescopes." (4 marks).
4. Write a sentence of your own with each of the following.
   (*a*) . . . so . . . that . . .
   (*b*) though. (6 marks).
5. Why may the stars never come near each other? (3 marks).

1969 Baccalaureate Part I Examination.
English Test (first foreign language).
Sections A and B.
Time: 2 hours. *Coefficient:* 2.

Read the following passage carefully and then answer the questions:

OUR CHANGING WORLD

The progress of science is never-ending. Every day brings new inventions and discoveries that make tomorrow just a little different from today. What the world will be like twenty-five years from now, no one can safely predict. We know that we are just beginning an air age — a period in which we shall need to become air-minded. Whether we ever fly an airplane or not, it will be worth-while to understand the principles by which air-crafts operate. Aviation and numerous other fields of endeavour are calling for young men well grounded in science to participate in the progress of the future. Opportunities in the field of new invention were never so great. Also great is the demand for persons to carry on work already under way.

BOWER and ROBINSON

Questions:

1. From the passage pick out 4 words having main stress on the second syllable. (2 marks).
2. Find in the passage the synonyms of: TO FORECAST; STARTING; TO TAKE PART IN; DEVELOPMENT. (2 marks).
3. From the passage pick out 2 compound adjectives and say how they are formed. (2 marks).
4. Re-write the following sentence without using the inversion of subject and verb: "ALSO GREAT IS THE DEMAND FOR PERSONS TO CARRY ON WORK ALREADY UNDER WAY." (1 mark).
5. Change to the Passive Voice: "Aviation and numerous other . . . progress of the future." (1 mark).
6. (*a*) Why can't anybody safely predict what the world will be like twenty-five years from now?
   (*b*) What is the great need in the field of science now? (4 marks).
7. Translate the first 4 sentences into Vietnamese. (8 marks).

1968 Baccalaureate Part II Examination.
English Test (second foreign language).
Section C
Time: 2 hours. *Coefficient:* 2.

Read the following passage carefully and then answer the questions:

IN A GARDEN

The fog was rising, lifting to the tops of the trees. I could see the woods at the end of the lawns. Above my head a pale sun tried to penetrate the heavy sky. It was hotter than ever. A bee hummed by me in search of scent, bumbling, noisy, and then creeping inside a flower was suddenly silent. On the grass banks above the lawns the gardener started his mowing-machine. A startled linnet fled from the whirring blade toward the rose-garden. The gardener bent to the handles of the machine and walked slowly along the bank scattering the short-tipped grass and the pin-point daisyheads. The smell of the sweet warm grass came towards me on the air, and the sun shone down upon me full and strong from out of the white mist. I whistled for Jasper but he did not come. Perhaps he had followed Maxim when he went down to the beach.

DAPHNE DU MAURIER *Rebecca*

Notes: *Linnet:* a small brown singing bird
Whir : sound made by machine in rapid motion.

Questions:

1. Ask questions to which the underlined words are the answers:
   (*a*) *The fog* was rising.
   (*b*) I could see *the woods* at the end of the lawns.
   (*c*) A startled linnet fled from *the whirring blade*.
   (*d*) *The gardener* started his mowing-machine. (2 marks).
2. Put in the interrogative form:
   The fog was rising. I could see the woods. (1 mark).
3. Pick out: (*a*) 4 words which sound as in "*my*";
   (*b*) 4 others which sound as in "*cut*". (2 marks).
4. Write an adjective from: flower, sun;
   a noun from : warm, strong. (2 marks).
5. What was the bee doing? Why? (2 marks).
6. Why was the bee silent? (2 marks).
7. What is a lawn? Give another word for "To cut the lawn." (2 marks).
8. Is there any difference between "grass" and "weed"?
   Write a sentence with each word. (2 marks).
9. What is a daisy? Name 5 more things of the same kind. (2 marks).
10. *Briefly* describe an afternoon spent in a garden or in a park. (3 marks).

Discussion

Malaysia: The objectives of English language teaching in Malaysia have not been precisely spelled out. In general the objective is stated to be complete mastery of the four language skills; i.e., a high standard of achievement is expected. The syllabus for English teaching specifies course content rather than learning objectives, so that the examinations, developed and administered through a national central examination syndicate, actually constitute specifications of the objectives.

We may consider the examination administered at the end of the third year of secondary school — the Lower Certificate of Education examination — as an illustration of the type of examination in use in Malaysia. There are actually two different examinations used at this level: one for schools where English is the medium of instruction, the other for schools where Malay is the medium. In the English-medium schools, the examination consists of the following:

1. A test of written expression, in the form of a free composition of approximately 300 words on any topic within the child's range of experience. No guidance as to format, topic, etc., is given to the student for this composition; 1½ hours is allotted.
2. A test of reading comprehension of a passage of 250–300 words, the student being required to answer questions on the passage in English.
3. A vocabulary test, generally on the meaning of words used in the reading comprehension selection.
4. A test of the student's ability to summarize the main idea of a reading passage.
5. A test on usage of various parts of speech.

For Malay-medium schools the test is similar, but shorter and simplified. One hour is allotted for the composition, which is expected to be of approximately 120 words. The reading comprehension passage is of 150–200 words and the questions set are somewhat simpler. There are also a vocabulary test (words in context) and a test of grammar usage; no summarizing is asked for.

Neither examination includes any testing of listening comprehension or oral production.

It should be noted that as of this year, Malaysia is changing over entirely to Malay-medium instruction, beginning in the First Primary and adding one grade level each year, so that eventually Malay will be the medium of instruction at all levels in all schools. English is to become the compulsory second language for all students, with the objective of achieving English-Malay bilingualism by the end of schooling (14 years). English is considered to be of special importance for scientific and technical uses.

Singapore: To illustrate the English examination system in Singapore, we can focus on the examinations set at the end of the fourth year of secondary in the non-English-medium schools. First, however, the language situation in Singapore must be explained. The country has four official languages: Malay, English, Chinese (Mandarin), and Tamil. Each of these languages is used as a medium of instruction, so that we have English-medium, Malay-medium, Chinese-medium, and Tamil-medium schools. In non-English-medium schools English is taught as the second language; in English-medium schools one of the other official languages is taught as second language. The objective is the achievement of bi-lingualism, English plus one other of the official languages.

Malay has the status of "the national language", but English is the dominant language of internal and external communication, the language of government administration, courts of law, business, and most higher education. English is the *lingua franca* of the educated, and is a required "passport" for obtaining most jobs in Singapore. The emphasis in English teaching is on the use of the language for communication, and oral skills are emphasized.

At the end of Secondary Four, students of the Chinese and Tamil streams sit for an examination which includes tests of written expression (free composition) and tests of comprehension and usage. This examination, however, is to be discontinued after this year, to be replaced by the new Singapore-Cambridge Certificate Examination.

Students in the Malay-medium schools have a choice of examinations. They may sit either for the "special paper" which Singapore inherited from the time preceding its separation from Malaysia, or they may sit for the same exam given to the Chinese and Tamil-medium students. Since the special paper is given only a Pass-Fail grade, whereas the other exams can be rated Fail, Pass or Distinction, the superior students of the Malay stream are encouraged to take the regular examination with its possibility of a "Distinction" rating.

In the examinations at this level there is no test of oral skills. These, however, are emphasized in teaching, and are included in the examinations at the end of Primary (6th year) and at the 12th year level.

Thailand: English is taught in Thailand as a foreign language beginning in the fifth year of primary and continuing during eight years of schooling. There are no accurate statistics on the role of the English language in Thai life, but its utility is generally obvious. There are two chief motivations for learning English; viz., its economic value for securing employment, conducting business, etc., and the requirement for English for higher studies. The second is probably the strongest motivation, and the English teaching program in general is geared to the utilization of English in higher education.

The examinations given at the end of secondary level are prepared by the Ministry of Education. They are based on the national syllabus for English and function both as termination exams for secondary and as tests for university entrance. The students are tested chiefly in reading, writing, and grammar. Testing of oral skills is notably absent from these examinations; there seems at present no way to overcome the administrative problems of administering satisfactory oral examinations to the large numbers of students nation-wide.

The examinations include a test of reading which requires the student to answer questions on set reading passages; a vocabulary test on words in reading context; a writing test in the form of semi-free guided composition; and an objective test on grammatical structures.

The examinations taken by the individual student depend on his special area of concentration in secondary school. There are three such areas: arts, sciences, and general. The examinations include three sections. The first is required of all students of all three areas of concentration. The second is required of arts students and science students, but not of general program students; it consists especially of questions on assigned readings. The third section is exclusively for arts students, and tests grammar and reading comprehension in more depth than do the other sections.

Indonesia: In Indonesia, English is the compulsory "first foreign language" beginning at secondary level; in effect, it is generally the third language for the student, who speaks at home one of the many languages native to the country (Javanese, Sundanese, Minangkabau, Buginese, etc.), and learns in school the national language, Bahasa Indonesia, which is the medium of instruction. The goals of English instruction are a "working knowledge of English" at the end of secondary, with priority given to reading skills.

The examinations at this level are external examinations administered at province level, and test reading comprehension, vocabulary and grammar; there are no external examinations of oral proficiency. The results of these examinations are often very disappointing, and reflect the difficulties of the current English teaching situation. The program suffers, for instance, from serious shortages of textbooks, especially outside Java. A set of texts designed for lower secondary English classes (the "Ministry Materials") was prepared in 1962, but funds for printing are inadequate, and there are problems of distribution (e.g., 150,000 textbooks have been in storage in Djakarta since Dec., 1969, because no way has yet been found to ship them to the schools they are designated for). For upper secondary level, textbooks are in preparation but will not be ready for two years. Meanwhile the teachers do whatever they can. Teacher inadequacy is also a problem; in West Java, for instance, only about ten per cent of the English teachers can be considered properly qualified. A teacher upgrading program is in progress, but its effect is only slowly being felt.

In the light of this, it is not surprising that the English language testing program must in general be considered deficient.

Vietnam: English in Viet Nam is not a second language, but a first or second foreign language. A student electing English as first foreign language takes it for seven years in secondary. The first national examination is administered at the end of the sixth year (First Baccalaureat). The student who passes this exam can proceed to the seventh year and sit for the Second Baccalaureat exam.

English in Viet Nam is considered part of the program of general education. Instruction is based on a national syllabus. The national examination test items are selected by the Ministry of Education from items proposed by the teachers.

Laos: English is considered the first foreign language in Laos, and as part of the General Education program. There is as yet no official statement of program objectives, and a national syllabus is currently being prepared. It is generally assumed by the teachers, most of whom are native speakers of English from abroad, that English should be studied in Laos as a tool for international and regional communication, for higher and technical education, and for increased economic opportunity. The English exam at the end of secondary is optional; it is an oral examination. The student is given a passage of 20–25 lines to read silently. He is then asked to read the passage aloud, answer some questions on it to demonstrate comprehension, then translate some sentences of the passage from English to French. Obviously the student's French is being tested as well as his English.

The chairman noted that, from a review of the panel presentations, the objectives of English language learning/teaching in the region could be enumerated as follows:

- bilingualism
- internal communication
- external communication
- government administration
- employment, economic opportunity
- vehicle for higher education
- vehicle for technical education
- vehicle for general education

Not all of these are applicable for all countries of the region. Also, the difficult task of translating objectives into precise statements of skill and levels of skill mastery required to reach the objectives, and the further task of developing teaching programs to meet these skills objectives, are still not complete.

In the discussion the panel was asked what each country's position was on national, centralized examinations. Mrs. Molly Ng Kuen Seong (Singapore) stated that, in Singapore, national examinations were a necessary tool for student selection and for maintaining uniform standards. Although Singapore educators deplore many of the adverse effects of the highly competitive examination system — e.g. student anxiety over passing exams, teaching aimed at exam-passing rather than genuine development — it is nonetheless necessary to have national examinations at the end of primary and secondary to screen students for the next highest level, because space and facilities are limited. Mr. Murad bin Mohd. Nor (Malaysia) stated that the same situation prevailed in Malaysia.

Mr. Banlue Tinpanga (Thailand) noted that in Thailand the attitude toward national examination was neutral, but that there were at present too many administrative problems — preparation, distribution, security, time — to make a national examination program feasible. Mr. Tirtopramono (Indonesia) agreed that for Indonesia the same statement applied, probably with even stronger force. Given the practical impossibility of successfully conducting national exams because of problems of cost, distribution and security, the discussion pro or con national examinations was only academic. However, even were national examinations feasible, the trend of thinking in the country seemed to be against national exams.

Dr. Leocadio (Philippines) stated that a national examination system had been tried in the Philippines, but was discontinued because it was found undesirable. It required considerable time, money, and administrative effort, but offered no results significant enough to counterbalance the restrictive effect it had on teaching and learning. It was considered better to allow schools flexibility to develop along individual lines; rather than set up a competition to pass exams, it was considered healthier to promote competition in other areas.

Dr. Duong Thanh Binh (Viet Nam) stated that although she was, in principle, neutral on the subject, there were good practical reasons for having a national system of examinations in Viet Nam. It was at the moment the only way that students could be given proper credit for their work, and also was a means to force the upgrading of weak teaching programmes and eliminate the defective programmes. However, she noted, the system had led to excessive examination-oriented study, especially among male students.

The chairman noted another advantage of a central nationalized examination and syllabus; viz., that innovation could be put quickly into effect throughout a country.

CONTEXT FOR LANGUAGE TESTING

by JOHN A. UPSHUR

Let me first say how happy I am to be here in this city and among this distinguished group of educators. Three months ago I did not expect that I would be able to attend this seminar on language testing. But a few weeks later I learned that I would be able to come[1] and at that time I was given a weekend to provide a title for the talk I'm giving today. I'm quite certain that there are people who can think of intriguing titles and then prepare remarks to fit, but this is a talent I lack.

My task, then, in selecting a title was to be as unrevealing and ambiguous as possible. In the limited time available, "Context for Language Testing" was the best I could manage. I thought this would allow me the choice of talking about the historical development of second language testing, technological aids in testing, the populations who are tested, the places and institutions using language tests, the reasons for testing, the professional qualifications for those who construct tests, views of language and language use which form a base for tests, questions of test validity and reliability, objectivity and compromise, or any of a number of other topics all important for one whose interest is language testing.

While I was still trying to choose a topic for my title, I heard at a recent conference on second language teaching, a talk which began and ended with the same sentence, "Language like every other blessing derives its value from its use alone."[2] The theme of that paper was that we can never fully understand the nature of language until we consider its use; the fundamental fact of language is that it is a tool for communicating something to somebody. I am convinced that the same can be said of testing. We are here because we recognise the value of testing; we should also realize that a test like every other blessing derives its value from its use alone. We can never fully understand the nature of testing until we consider its use; the fundamental fact of testing is that it is a tool for communicating something to somebody. The context for language testing that I will discuss is the communication environment of teaching and learning a second language in an educational program.

An educational program can be viewed as an information processing system. There are certainly many ways of conceptualizing educational programs — or anything else. All these ways of analyzing are simplifications. An educational program is so complex that we can't think about it in all its rich detail without constantly losing sight of basic parts. Some kind of simplification is, therefore, necessary. I've decided to look at an educational program as an information processing system, because I think this view will best allow us to focus upon the place of testing in education.

I should at this point make clear what I mean by an information processing system. By *system* I refer to a set of related or connected parts which form some kind of functional unity. We are accustomed to talking about languages as systems although some speak of form classes which occur in "slots" of sentence patterns, while others speak of "strings" related to other "strings" by "transformational rules".

The game of pocket billiards (or pool) can also be thought of as a system in which there are players, cue sticks, balls and a table with pockets. One player strikes the cue ball with his cue stick setting it in motion. It strikes another ball which moves in a direction determined by the angle at which the cue ball strikes it. That ball may in turn strike still other balls to set them in motion. Pocket billiards can be seen, therefore, as a mechanical system in which the parts — the balls — affect one another by the forces they apply to each other. In an information processing system the parts interact by giving and getting information rather than mechanical force.

Last December my wife and I decided to leave the ice and snow of the Michigan winter for a two week vacation on a tropical beach. En route we stopped for a few days in a city where I became quite suddenly aware of sensations all too common to travellers abroad. I said to my wife, "I don't feel very well. I think I've caught it."

She went to the hotel lobby and explained the situation to the clerk who called a doctor from a nearby clinic. He came to my room within the hour, gave me an injection, some pills and some reassuring words. By afternoon I felt fine and was on the plane headed for the beach.

I haven't recounted this episode to gain sympathy, but only to show how parts of a system can interact by giving and getting information. My wife's behavior depended upon the information she got from me. The clerk gave information to the doctor because he got information from my wife. Because communication in the system was good, the doctor arrived, I got well and enjoyed my vacation.

Let me now show an educational program as an information processing system. I've chosen the flow chart as a means of representation because it is easy to follow visually, and it allows one to talk in more or less detail. It is especially well suited to discussion of systems; and as I hope to demonstrate, tests are necessary parts of dynamic information processing systems.

A flow chart has three figures: rectangles, diamonds and trapezoids which are connected to one another by arrows. The rectangles contain statements about procedures to be followed; the diamonds contain questions and represent the need for making decisions; the trapezoids contain initial or final statements; and the arrows indicate the hierarchical or chronological relationships among the statements and questions.

Figure I shows a rather common and extremely simple educational program. In this system the student enters, we give him instruction and he leaves. We don't even ask how much he has learned. At least, if we ask, it doesn't affect the way the system works.

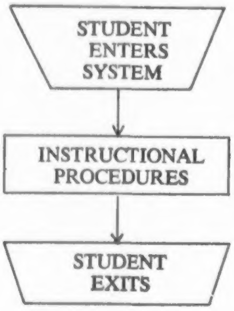


Fig. 1 A Very Simple Educational Program

Figure 2 is a flow chart of an educational program in which one continues to provide instruction for a student until he meets the criteria for success in the program. That is we continue to teach him until he learns.

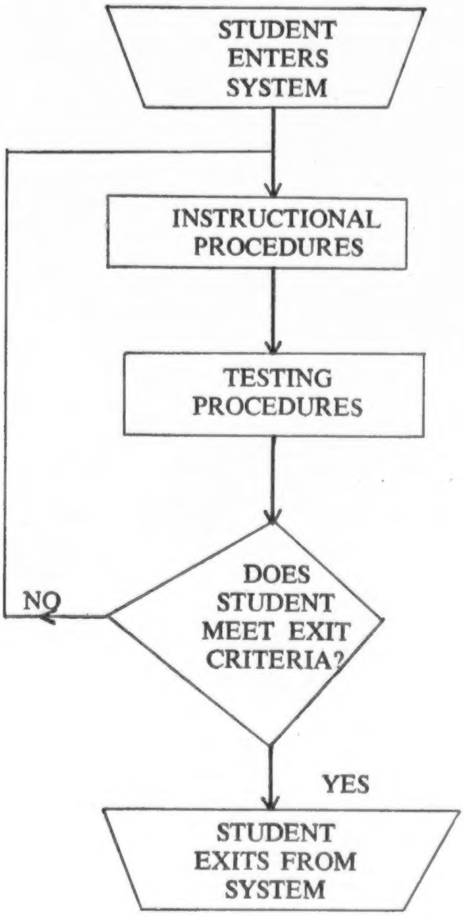

STUDENT ENTERS SYSTEM
INSTRUCTIONAL PROCEDURES
TESTING PROCEDURES
DOES STUDENT MEET EXIT CRITERIA?
NO
YES
STUDENT EXITS FROM SYSTEM

Fig. 2 A Simple Educational System

The student enters the course; we instruct him (in English as a second language for example); then we test him according to the goals of the course; we next ask whether he has met our performance criteria; if he has, we graduate him; if not we send him back for further instruction and continue as the arrows indicate.

Before I go on, I want to make an observation. The program of Figure 2 includes a question, it requires someone to make a decision. Information is needed in order to make the decision, and a test is the means for communicating information about the student to the decision-maker. In an educational program, or in any other system, tests will generally precede any anticipated decision points.

The program of Figure 2 has two serious flaws in its design. We teach any student who comes to us regardless of whether he needs instruction or not, and we have no way of getting rid of him if he does not learn. Figure 3 shows a somewhat more sophisticated program in which we test to see whether a student is qualified, that is whether he will learn, and we test him also in order to determine whether he already knows those things we are planning to teach him.

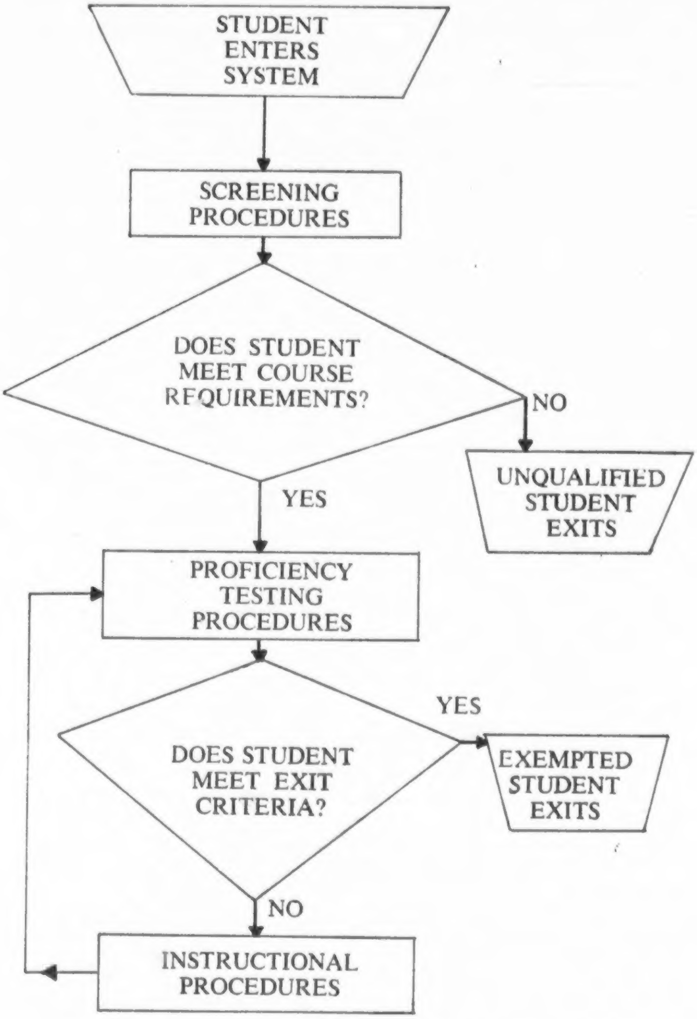

STUDENT ENTERS SYSTEM
SCREENING PROCEDURES
DOES STUDENT MEET COURSE REQUIREMENTS?
NO
UNQUALIFIED STUDENT EXITS
YES
PROFICIENCY TESTING PROCEDURES
YES
DOES STUDENT MEET EXIT CRITERIA?
EXEMPTED STUDENT EXITS
NO
INSTRUCTIONAL PROCEDURES

Fig. 3 Use of Information in an Education Program

I'm quite sure that the three programs I've just described do exist somewhere, but I want to describe to you a language teaching program I know of, which includes rather extensive, formal testing procedures. It is located at a reputable American educational institution. I won't name the school, however, because it probably is not unique.

This institution offers a three-year sequence in a modern foreign language. The first two years are primarily audio-lingual, the third is essentially a reading course. It is immaterial that the language is not English; it is an L^2 in an L^1 setting, just as English language instruction is in Thailand, for example.

Any student who wishes to study the language is admitted to the first-year course. At the end of the first year the students take a test based upon their year's work. The best 50 per cent of the students according to this test are admitted to the second-year course. Another test is administered at the end of the second year. This highly objective test is designed to measure mastery of second-year material, and like the first-year test yields a wide range of scores. Half of these students are required to end their study at this point; the students with the better test scores are admitted to the third year course — a reading course in which students select the works in the foreign language that they wish to read. These highly selected third year students all receive very good grades for this course.

The two major tests used in this program are soundly constructed. Each year's syllabus was first analyzed by a test construction team. This analysis was reviewed by the teaching staff and then appropriately modified. For each vocabulary list, grammatical construction, and so forth a set of five items was prepared. A group of teachers reviewed these items and selected the best three or four from each set for pre-testing. The pre-test results were analyzed and one item was taken from each set for the final test form. Selection of those items was made according to standard procedures in order to yield an instrument which would give the widest possible distribution of scores.

I have given only an outline of the considerable painstaking effort that went into test construction for this language program and perhaps one might conclude that here is an example of a highly effective testing program. If, however, we evaluate this program in the context of an information processing system, I think we will have to conclude that it is grossly inefficient, and, furthermore, it tends to inhibit good teaching in the course and it impedes the development of a better language teaching and learning program.

First of all, what is the primary use of test information for the system? The flow chart (Fig. 4) makes it obvious. The students enter the first course, are instructed and then tested. Then each student's score is looked at to see if he has been a high achiever or not. The high achievers are sent to the second year course. The low achievers are sent on their way elsewhere. The same procedure is followed for the second year.

The test procedure does give information about a student's attainment, but when we consider its *use*, we note that the information is used only to determine a student's aptitude for further study. How much better it would have been to administer a foreign language aptitude test at the start of the course and not have students waste a year or even two years in the course before they are rejected. Perhaps it is not so bad to use a student's ability to learn in a first-year audio-lingual class to predict his ability to learn in a second-year audio-lingual class. Less defensible is using a measure of a student's ability in an audio-lingual course to determine his ability to *read* in a second language; yet this is the basis for selection of third year students in this program.

I made two other objections to the use of tests in this program; that they impede the development of a better program and that they inhibit good teaching.

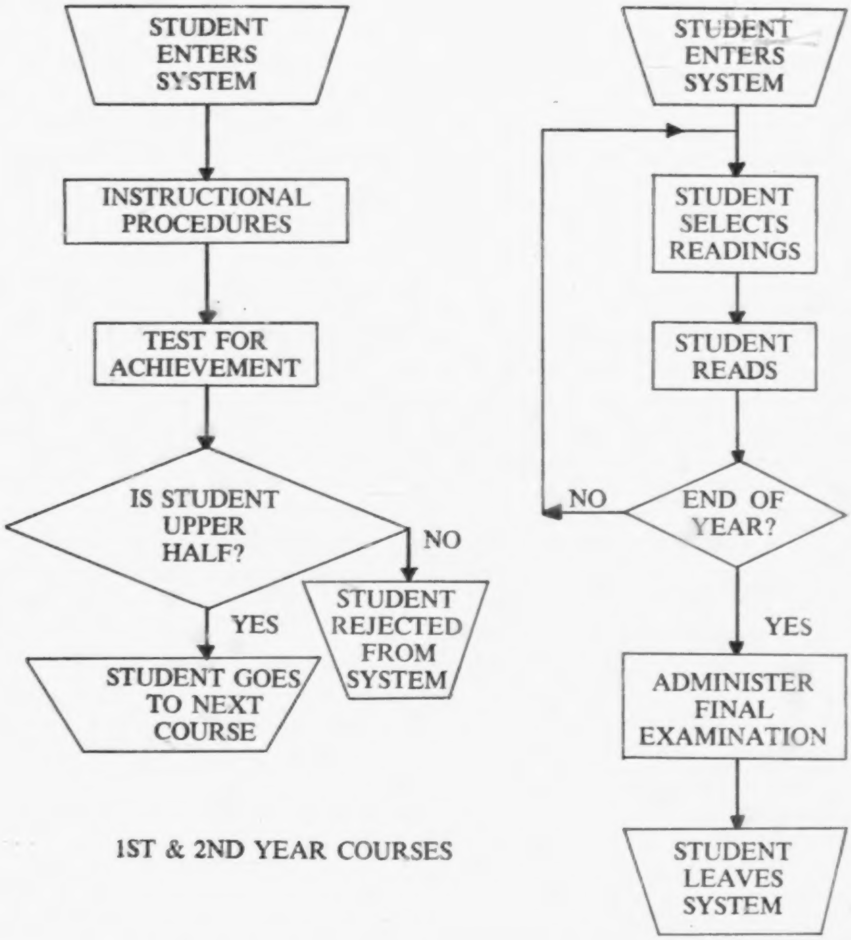

STUDENT ENTERS SYSTEM
INSTRUCTIONAL PROCEDURES
TEST FOR ACHIEVEMENT
IS STUDENT UPPER HALF?
NO
YES
STUDENT REJECTED FROM SYSTEM
STUDENT GOES TO NEXT COURSE
1ST & 2ND YEAR COURSES
STUDENT ENTERS SYSTEM
STUDENT SELECTS READINGS
STUDENT READS
END OF YEAR?
NO
YES
ADMINISTER FINAL EXAMINATION
STUDENT LEAVES SYSTEM

Fig. 4 Use of Final Examination Information

Improvement of a program implies change in the program. Changes result from decisions, and people require information in order to reach decisions. But the system I have described does not provide any means for communicating to anyone the information which will enable him to decide on changes. The formal requirements of the system are only that in each of the first two years half of the students are rejected from the system, and that the remainder of the students leave the system at the end of the third year. The operation of the system is in no way altered in response to the students who enter the program.

The formal tests are used in this program to identify upper-half and lower-half students in terms of their year's achievements. Whereas a test can make these identifications quite reliably in a group which contains students who are either genuinely good or very poor, the identification is much less reliable when achievement is more normally distributed: a few very good students, a few poor ones and many who are average. For a teacher to produce a class of either good or poor students, his best strategy is to teach half the students and ignore the other half. Teachers in this system are conscious of the eventual tests and their use, and they are unconsciously inclined to teach only half their students so that test results will be clear cut. They are inclined also to identify the "good" and "bad" as early as possible. In this way the instructional component of the program tends to parallel the operation of the larger system of which it is a part. Early in the course the teacher performs the same function that the entire course will ultimately perform, that is, identifying the "good" and the "bad" and rejecting the latter. At this point, the "bad" are not formally removed from the program; the teacher just ignores them. Creative teaching is unnecessary since the goal of the program is already accomplished. Almost anything the teacher does will result in the "good" achieving more in the year than the "bad" do.

Here we have a double example of the self-fulfilling prophecy: the teacher's early judgements are confirmed by the final test results; and the program's designers are pleased to have confirmation of their hypothesis that only half the students in any year learn enough to continue. The system has been justified.

I have tried to show how a poor system can tend to produce poor teaching. But we know that good teaching does exist. Can this good teaching be used as a model for good systems, for good educational programs? I think that it can; and I think further that the way it can do so is by showing how and when to test.

Figure 5 is a generalized flow chart of a teaching-learning situation in which the teacher has already decided (or learned) what to teach and in which some kind of feedback from the teacher is necessary for learning. In essence, this system consists of two overlapping tests.

The problem posed by the teacher may be as simple as, "Repeat after me: /hæt/." He listens to the student response; was it /hat/ or /hæt/?, and judges the student's ability to hear and produce a particular English vowel contrast. This is one test.

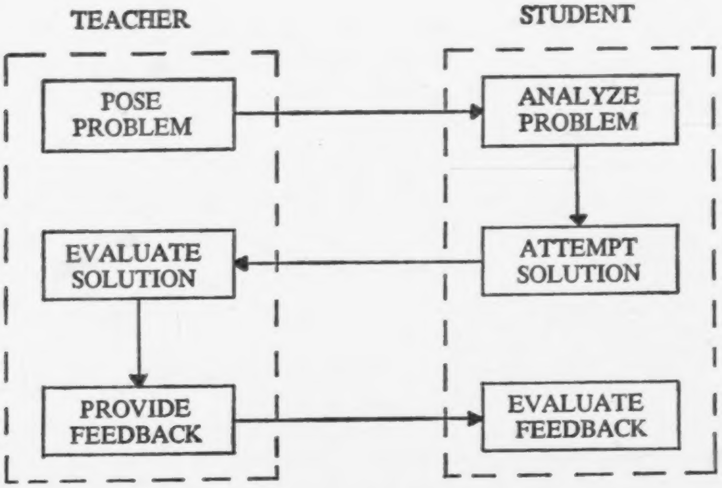


Fig. 5 Teaching-Learning Model

The student says /hæt/ or /hat/ in order to have the teacher say "OK," or "No, try it again." This is the other test. In the teaching-learning situation, the teacher uses test information in order to know what kind of feedback to provide. The student uses his test information in order to know whether his performance is acceptable, or whether it must be altered in some way.

This simple scheme brings into focus two related facts: the teacher knows what is "correct" or "appropriate" whereas the student does not yet know; the related fact is that the student is using test information to change his own behavior, but the teacher is not changing. In summary, the teaching-learning situation is one in which two people are testing each other in order that one of them can change: the one who changes we have termed "student".

I suggested earlier that good teaching does exist and it can be taken as a model for educational and testing programs. But we know that good teachers are constantly learning and improving; the good teacher is most often characterized as creative. Learning, improving, creating, all imply change. Thus the good teacher is also a "student". This is illustrated in Figure 6. Here the student begins by posing the problem. His presence in class says, in effect, "Teach me." The teacher analyzes that instructional problem and attempts a solution. He then observes the student's response to his instruction. If it is not satisfactory he changes his analysis of the teaching problem or his attempted solution. He has become the student.

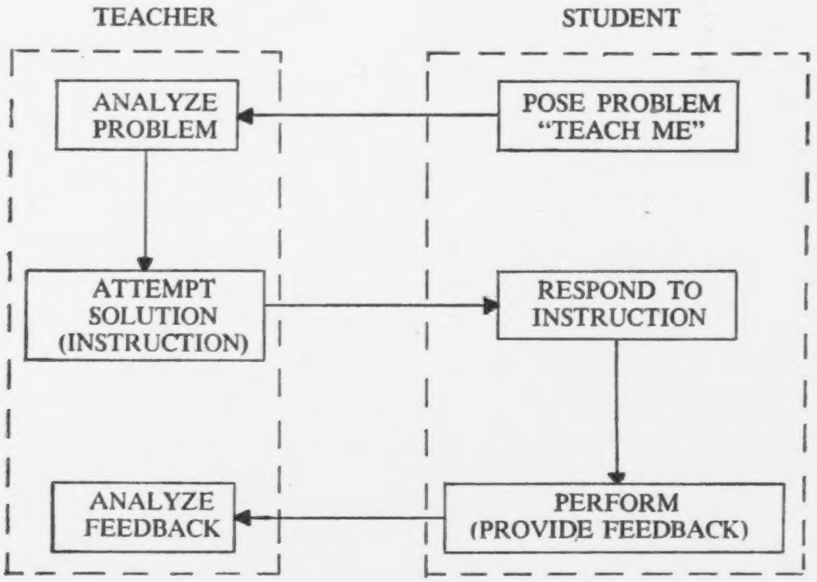


Fig. 6 The Teacher As Student

The good teaching model is, therefore, a *dynamic* information processing system in which each participant tests the other in order that the teacher may learn what and how to teach, and the student may learn that which is being taught. In Figure 7 I've attempted to indicate some of the dynamics of the system.

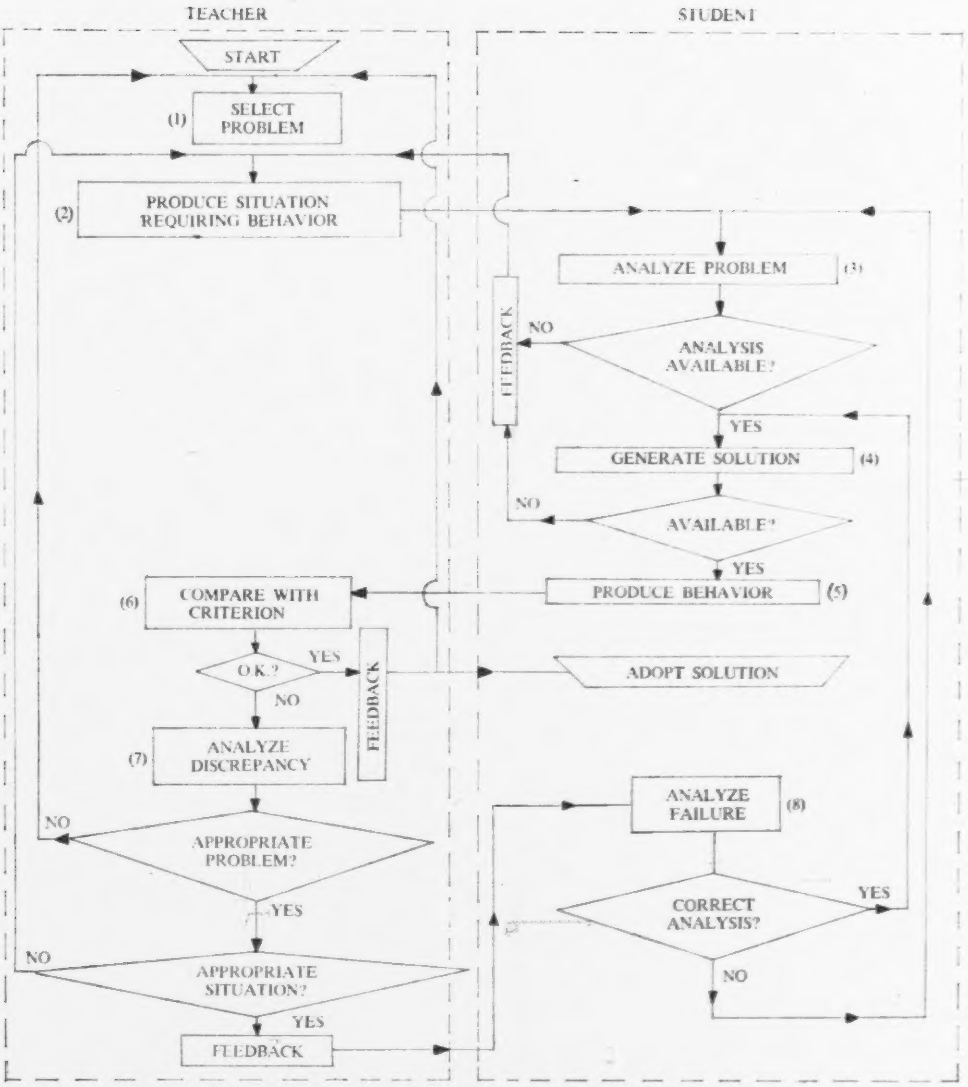

TEACHER
STUDENT
START
(1)
SELECT PROBLEM
(2)
PRODUCE SITUATION REQUIRING BEHAVIOR
ANALYZE PROBLEM
(3)
FEEDBACK
NO
ANALYSIS AVAILABLE?
YES
GENERATE SOLUTION
(4)
NO
AVAILABLE?
YES
PRODUCE BEHAVIOR
(5)
(6)
COMPARE WITH CRITERION
O.K.?
YES
NO
FEEDBACK
ADOPT SOLUTION
(7)
ANALYZE DISCREPANCY
ANALYZE FAILURE
(8)
NO
APPROPRIATE PROBLEM?
YES
CORRECT ANALYSIS?
YES
NO
NO
APPROPRIATE SITUATION?
YES
FEEDBACK

Fig. 7 Dynamic Teaching-Learning (Testing) System

When he begins, the teacher's first task is to select what is to be learned (1), and present an appropriate task for the student (2). The student then attempts to analyze the problem in order to determine what is required of him (3), and then to devise a way of doing it (4). If he fails to understand the problem or to see a possible solution, he may inform the teacher who will then produce a more appropriate task (2 again). If the student is able to understand the problem and generate a solution, he produces some behaviour (5) which the teacher can observe and compare with a criterion for adequate performance (6). If the student's performance is satisfactory, the teacher lets him know. The student accepts this as confirmation of his solution; he adopts the hypothesis by which he derived the solution. If, however, his performance is not satisfactory the teacher must try to decide what went wrong (7). He may decide to select a simpler problem to be learned (1), or to provide a different situation for the same problem (2); on the other hand, he may signal the student to try again. In that case, the student must analyze his failure (8) and either attempt a new solution (4) or a new analysis of the problem (3) depending upon why he thinks he failed.

In this system testing is constant. The language student tests in order to discover discrepancies between the hypotheses he generates and the rules which govern correct behavior. At the same time he attempts to minimize these discrepancies. In this way he learns the new language. The teacher tests in order to discover discrepancies between student behavior and the criterion. At the same time he works to minimize these discrepancies. In this way he improves his teaching.

Test information is used in this system for deciding what to teach, how to teach and whether learning has occurred. Test information can also be used to decide on the adequacy of the criteria for learning. (This use is omitted from figure 7 in order to keep the flow chart from becoming too complex and difficult to follow. A circuit from (7) to (6) is easy to add, however.)

In the same way that a good teacher uses test information in the classroom, a good education program should provide means of communication which allow it to change. In this larger system also, tests provide information to guide changes in what is taught, how it is taught, and the criteria by which learning is recognized and defined. There is another important question for tests to answer. Because an educational program like every other blessing derives its value from its use alone, what is the value of the learning which does occur?

Decisions must be made if a dynamic system is to operate, and tests can be useful, and therefore valuable, when they provide appropriate information to those who will make the decisions.

One general method for reaching decisions is "passing the buck." In the case of the "what to teach" question, this method might take the form of asking an examiner from Cambridge what to teach, or asking a linguist from Canberra. This might be a satisfactory first step, but it is not sufficient for a dynamic system which is to change and improve. Another method is to specify the initial state of a problem and the target state in order to determine differences which must be reduced; the differences are the things to teach. Contrastive analysis is an example of this method. But contrastive analysis is not enough. In most English teaching programs students enter with some knowledge and skill in the language, and we don't want to waste time and effort teaching things that students already know. A test can tell us what these things are. A test can tell us also what the students want to know. We are never able to teach our students much that they don't wish to learn. The test itself can be quite simple: ask the students what they want — and ask again next week and the week after, because just as their knowledge can change, so can their aspirations. There is another test which will help us decide what to teach. In a program which is constrained by time, we cannot expect to teach everything, but we *can* expect to teach more to some students than to others. A language aptitude test can help us to decide what to teach to whom.

I should mention one test which is *not* taken by the students. We noted earlier that the teacher is defined as the one who knows what the student is to learn. Deciding on what to teach must depend upon what the teacher knows. Testing the teacher can give this information. I realize that I am offering an unpopular suggestion; but remember that we are considering dynamic systems. Teachers also change. They can learn things which can then be included among those things that students are to learn.

Let me offer a suggestion also for answering the question of how to teach. A useful method is to decide on techniques or operations which have proved effective in "similar" situations in the past. How do we know when techniques are successful? Tests give us measures of success. But how do we decide when situations are "similar"? There is certainly similarity between two situations in which students are comparable with respect to what they already know, what they want to learn and their aptitudes for learning; and we have already seen that testing can give us that information.

Tests can also give us information about reasonable criteria of learning. Statements about what to teach are often made in linguistic terms: for example, comprehension of questions of the form, "Isn't it in Bangkok that we're meeting?" A reply of "yes" can be interpreted as an indication of comprehension, but this reply is not adequate. If we are to conclude that the student has indeed learned, how much time do we give him to think, to decode, or to translate? Should he respond in five seconds? Three? One? Many battles have been waged over questions like this, but I wish to talk only about the use of tests in establishing criteria of learning.

In this example of comprehension teaching, a teacher may first accept from a student a reply of "yes" at any time after the question was posed. He can then adopt as a criterion of learning that the student respond more quickly than before. Accordingly, so long as the student increases his speed, learning is occurring. When the student's performance stabilizes, the teacher concludes that a criterion of greater speed would not be reasonable for the student, given the techniques and material available. At that point he goes on to teach something else.

Before continuing I want to make clear what a criterion is. First of all, it is not the same thing as a program goal. It is only a standard to which test behaviour can be compared. When behaviour will be judged as either "satisfactory" or "unsatisfactory" the criterion which will yield the most test information is the one which will be satisfactorily met about half of the time.

Just as tests can be used in the classroom to recognize a student's increasing ability to understand English, tests can be used to recognize an educational program's increasing capacity to teach English. Tests to determine what the program is able to teach at any point in time establish reasonable criteria for evaluating in-service teacher training, the addition of language laboratories, changes in instructional techniques or almost any other change which can be made in the program.

When we have decided what to teach, how to teach, and what the criteria of learning are, there are no problems in discovering when learning has occurred. A test to provide this information has been completely specified.

There is one more general method for making decisions that I will discuss. This is called "planning". In planning, a simplified description of a complex problem is constructed, and means are devised for solving that simplified problem. One hopes that its solution will result in the solution of the complex problem. The complex problem which English teaching programs attempt to solve is producing students who can communicate with English speakers by means of the English language. This complex problem is often simplified to teaching a knowledge of English. We have what is known as a linguistically oriented program when "knowledge of English" is characterized as the ability to decode and construct sentences which conform to the rules of English grammar as set forth by linguists.

It is certainly possible to solve simple problems while complex problems remain. Tests of achievement in our courses and of proficiency in communication can show us how valuable our programs are. When we find that students who attain more in the program do not become correspondingly better communicators, we must be prepared to revise our plans.

I've been talking about language testing and change. The context I chose is the system which includes within itself the capacity to grow and improve. Tests have value in this system because of their use; in a dynamic educational program we can do more for our students tomorrow than we were able to do today.

Discussion

Discussants: Upshur; Cherrier, Enriquez, Mowla, Otter

It was observed that the dynamic information processing model proposed by Prof. Upshur might be of special relevance to the South East Asian Area. The school systems in the area certainly have the capacity to grow and change, probably more so than in a country like the United States, where forms and patterns are already quite rigidly established. The South East Asian area is acutely conscious of change and development; the problem here is not rigidity, but lack of direction for change, i.e. a lack of information to serve as the basis for change.

Specifically it was observed that syllabus development could take this model into account. Many countries in the area were in the process either of revising English syllabi or developing them for the first time. Professor Davis had previously noted that a syllabus might be considered as a collection of all possible tests, and, as corollary, a test would be considered one choice among the possibilities. The syllabus could be viewed in this manner, and furthermore incorporated as part of a dynamic information processing systems model. This would offer new and exciting possibilities to syllabus designers.

---

1. This paper was made possible thanks to travel funds granted to the author by the University of Pittsburgh Bangkok Project.
2. John W. Oller, Jr. "Linguistics and the Pragmatics of Communication," paper delivered to the TESOL annual convention, San Francisco, California, March 20, 1970.

CONTRASTIVE ANALYSIS AND ENGLISH LANGUAGE TESTING

by GERHARD NICKEL

The field of error analysis lies, for several reasons, in the care of foreign language teaching. It is of relevance for the teacher in general as well as for the test specialist, the textbook author and for those responsible for the drawing up of curricula. The psychology of learning sees the pupil's aberration as a normal process, just as the physician regards a high temperature in a sense as a positive phenomenon enabling him to make a diagnosis. Consequently, the main theme of the present discussion will be the diagnostic value of the mistake.

A survey of the literature available on this problem of didactics reveals a clear division into two camps. On the one hand, there are those who consider mistakes to be solely a negative phenomenon. They ascribe them to a lack of diligence, intelligence, concentration, and comprehension, that is to an insufficient command of the learning material. This occasionally gives rise to reactions on the teachers' part which range from slight disappointment to deep resentment.

This attitude, which is altogether understandable in view of the fact that the teacher is overburdened with work and is left little time to deal with diagnostic aspects of teaching, is based on a simplified model of the learning process. Seen in the light of this model, the pupil's acquisition of the regularities of the foreign language takes place as a linear sequence of individual learning steps resulting in a command of the structures taught or in a failure to command them. Mistakes are an indication of the failure to master the material and immediately require remedial measures.

Modern research no longer shares this view. One assumes a dynamic cycle, which replaces the line by a circle, in the course of which feedback effects come into play enabling the learner to make further progress by means of a trial-and-error procedure and to use his own mistakes as points of orientation. In this case the mistake fulfils a corrective function, represents a necessary intermediate state, and thus serves as a point of departure for a new cycle of progress. There is no other way to explain the relatively rapid progress in the process of foreign-language learning.

Among those holding the latter view are scientists such as S. P. Corder, who regards mistakes as signs of a particular learning strategy which must be made to influence the teaching strategy[1]. A similar view is put forward by P. Strevens[2]. Finally one must mention V. J. Cook, who likewise interprets pupils' aberrations as attempts to prove the correctness or incorrectness of parts of a hypothesis which the learner sets up concerning the target language[3].

The standpoint of each group is closely related to the answer each gives to the following question: Does the acquisition of the second language take place in an analogous way to that of the mother tongue?

Although the group of those claiming a clear analogy is undoubtedly decreasing nowadays, it cannot be denied that a major part of what have become known as direct and audiovisual methods is based on the illusionary assumption that the two processes are more or less analogous to each other. On the other hand they have constantly failed to draw the consequences necessary for the interpretation of mistakes and for the linguistic progression of teaching material. It has been shown, for instance by R. Brown and U. Bellugi, that the mother tongue is not by any means acquired in such a way that in the course of a learning strategy made up of small steps only regularities of an increasing degree of difficulty are learnt; instead steps involving mistakes are systematically built in (cf. e.g., negation in English)[4] The preference for lexical units, one of the main reasons for which is probably their high information content, leads to the production of torso or reduced sentences, which are then expanded through the parents' influence into grammatically complete sentences[5].

A certainly much larger group holds a deviant view, according to which there exist significant differences, the main difference being simply due to the fact that one language system has already been stored and that the latter exerts both a positive and negative influence on other languages learnt. Here there are again deviating views as to the nature of and reasons for interference. According to some linguists, e.g. S. P. Corder, the reason for the use of elements of the mother tongue instead of those of the target language simply is the lack of knowledge of the new rules, that is to say they do not believe in dynamic overlapping processes[6]. According to others there are genuine mixing processes which may even result in a sort of "interim grammar"[7].

Even if the two language learning processes were only approximately alike this would automatically necessitate a different attitude towards deviations. The fact that linguists have abstained from adopting this attitude can be interpreted as evidence that either they have never seriously thought of equating the two language learning processes or that they have so idealized the process of learning the mother tongue that they were unable to recognize the significance of deviations in building up the mother tongue matrix.

The evaluation of contrastive language analysis depends on the attitude towards the central question mentioned above. C. C. Fries holds the view that effective teaching methods must always take into account the structure of the mother tongue[8].

> The most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner.

A similar view is put forward by R. Lado[9]:

> Textbooks should be graded as to grammatical structure, pronunciation, vocabulary, and cultural content. And grading can be done best after the kind of comparison we are presenting here.

Occasionally contrastive linguistics is accused of overestimating its own importance[10]. However, we have constantly pointed out that contrastive linguistics only represents one component among others in an approach to the problems of error analysis and to the preparation of teaching material[11]. It must be stressed here that contrastive linguistics is not directed against what has become known as the "direct method". On the contrary: since it admits that the native language has considerable influence on the learning of a new language it supports, by its very nature, an intensive use of the target language as a countermeasure to interference.

Of course, opponents of this branch of applied linguistics are not lacking. We quote just one example[12]:

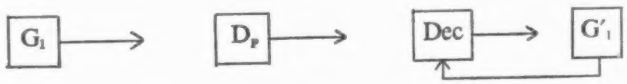
> The major contribution of the linguist to language teaching was seen as an intensive contrastive study of the systems of the second language and

> the mother-tongue of the learner. . . . Teachers have not always been very impressed by this contribution from the linguist for the reason that their practical experience has usually already shown them where these difficulties lie and they have not felt that the contribution of the linguist has provided them with any significantly new information. They noted for example that many of the errors with which they were familiar were not predicted by the linguist anyway.

Here it seems appropriate to state the following: it is undeniable that quite a number of teachers have for a long time recognized the difficulties that have their origin in the mother tongue. On the one hand, however, this is not true of all teachers; on the other hand, a systematic comparison of two language systems provides the opportunity for differentiating more exactly among the individual difficulties. Moreover, interference mistakes between the source and target language are, of course, not the only type of mistakes. In addition to those which occur when several foreign languages are being learned, there are intrastructural mistakes which are simply due to the fact that the new foreign language has not yet settled down in the pupil's memory and that he is mixing up rules within the foreign language. Finally there are "psychological mistakes" which are not predictable by any linguist, e.g. deviations due to the lack of concentration. In any case the claim put forward by two authors that teaching material drawn up on contrastive basis is detrimental to learning a foreign language seems to be contrary to all empirical observation, although here once again it must be stressed that in the case of the teaching material itself other factors, e.g. methodological ones, must be taken into account such as a meaningful contextualized representation of examples[13].

There is no doubt that contrastive analysis must be regarded as an important contribution of linguistics to the problem of foreign language teaching[14]. Symposia and research centres dealing with this problem are increasing in number. Monographs have emerged both from a round table meeting in Washington, D.C. and from the Second International Congress of Applied Linguistics in Cambridge[15]. The German Language Institute in Mannheim discussed problems of contrastive linguistics at a conference in Spring 1969 also[16].

Analyses of this kind aim at predicting learning difficulties, at explaining them, and at meeting them by means of adequately graded teaching material (qualitatively and quantitatively). Even more ambitious was the plan of some contrastive linguists to master the central problem of linguistic difficulties on the basis of contrastive analyses[17]. Most of these contrastive studies are based on a systematic comparison of the source and target language. In order to grasp the import of contrastive linguistics one has to visualize the first language learning process as follows:



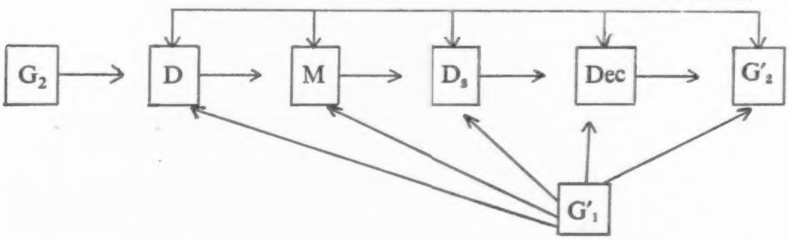
G_1 = grammar of the mother tongue
D_p = primary linguistic data
Dec = decoding
G'_1 = acquired grammar[18]

This communication process is neither ecoromical nor easy. Yet success is guaranteed by strong motivation, the amount of teaching effort involved (on the part of playmates, parents, etc.) and the usually clear relationships between the primary linguistic data and the situational context.

Applied to the learning of a second language this model only makes sense if the second language is learnt in the country in which it is the main language under conditions which are very similar to those mentioned in connection with the process of learning the mother tongue. Here, too, however, as in the case of bilingualism, it is certainly necessary to modify the model in some way. Moreover, even in the case of bilingualism we do not believe that the two linguistic systems have absolutely equal rank nor that anyone has a perfect command of two languages. The human brain seems to have difficulties in storing the data of individual language systems separately. This is perhaps because the systems are not completely rebuilt and stored separately; they are rather in some ways in contact with each other. This is a further argument in support of contrastive linguistics.

Since we wish to deal with the process of foreign language learning, which usually takes place in the area of culture of the source language and under classroom conditions, we have to deal with different circumstances. Now it is no longer primary data that are being presented, but secondary data which have been prepared in the light of didactics and methodology. Didactics refers to the manipulation of the data (the "what") while methodology refers largely to the psychological conditions of learning, i.e. to the "how" of the presentation of linguistic data. But even this enlarged model presupposes that there is no previous linguistic knowledge and neglects the fact that the child has already acquired one linguistic system, namely that of his mother tongue.

We all know that when learning a foreign language, conceptions belonging to the mother tongue encroach on the target language, the amount of encroachment varying from learner to learner. As conceived by applied contrastive linguistics, the model of second language acquisition under classroom conditions is as follows:



G'_1 = acquired grammar of mother tongue

G_2 = grammar of target language

D = didactics

M = methodology

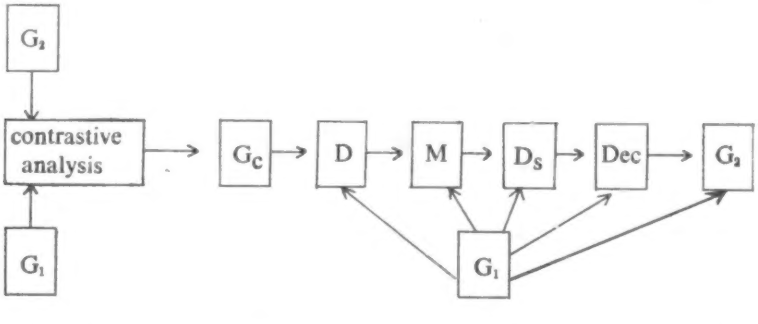
D_s = secondary linguistic data

Dec = decoding

G'_2 = acquired grammar of target language

Didactics, methodology and decoding are influenced by G'_1, and also by the grammar acquired in the target language (G'_2). This also shows that grading teaching material depends upon mother tongue conditions and also upon the knowledge acquired in the target language.

The aim of a contrastive analysis of two languages is to describe a partial grammar G_C, which consists of the sum of differences existing between the grammar of the native language G_1 and the grammar of the target language G_2.



This differential grammar resulting from a contrastive analysis is at the centre of the didactic programming, which of course does not mean that those parts of two linguistic systems which are in agreement with each other are not taught at all, especially since the overall linguistic system of the new language must be presented. At all events, selection and grading of the material are formed by it. Within the structure of the latter grading of the exercises and sequencing of the teaching material are influenced to the same extent.

Once again it must be clearly kept in mind that the overall composition of the teaching material is by no means exclusively determined by the results of the contrastive analysis. Modern language teaching must take account of all current developments in all linguistic fields, a requirement which is more especially fulfilled in the fairly natural situation of dialogue. On such an occasion one cannot avoid falling back on particular linguistic expressions despite their considerable difficulty, simply on account of their high frequency ratio. One need only think of the so-called irregular verbs. Time and again it will be necessary to work out compromises which take into account various factors, as has already been mentioned above.

The Project on Applied Contrastive Linguistics (*Projekt für Angewandte Kontrastive Sprachwissenschaft = PAKS*), which is being carried on at the University of Stuttgart, has in the main been devised to pursue the objectives listed above[20], as can be seen from previous PAKS publications. In view of the practical objective of PAKS, especially in connection with the problem of learning difficulties, it seemed necessary to undertake an empirical examination of the results of contrastive analysis. From practical experience it is known that degrees of difficulty in the learning process do not depend only on contrastive conditions. Error analysis must not be seen merely as a sort of procedure for checking contrastive analysis, but rather as a complementary which tries to take a short cut. One can arrive at error analysis via contrastive analysis or vice versa. However, the two approaches should not be regarded as equivalent since there are a fair number of mistakes which are not due to an interference between the mother tongue and the foreign language and which therefore cannot be dealt with in a contrastive analysis in the narrower sense[21]. In any case it should be clear that without contrastive linguistics all genuine error analysis is impossible if by "genuine" we mean that it is to achieve the higher goal of observing the real learning process which is strongly influenced by the native language.

Our project "Lapsology" is divided into three main branches[22]:

(*a*) *Descriptive lapsology:* Here we are concerned with the exact occurrence of the errors and with the reasons for the aberrations. From the practical point of view we do not think it advisable to separate the *why* and the *where*, that is to say in describing linguistic errors we do not take into account the native language, since language teaching is always concerned with devising immediate corrective measures which is not possible without an analysis of the error sources.

Errors are determined by the type of test involved and by the linguistic level. Between the two factors there is a clear relation of dependence. So in a dictation one hardly finds any syntactical mistakes. In this connection it is interesting to note that translations (about the didactic value of which the last word has probably by no means been spoken) contain fewer syntactical errors than, e.g. reproductions. The explanation lies perhaps in the fact that in doing a translation the pupil focuses his attention from the outset on grammatical difficulties, whereas in the case of reproduction or composition his attention is mainly directed to the content.

In dealing with interference phenomena we distinguish between interstructural and intrastructural ones, structure referring here to the total structure of a language. Of course, one can also use the forms "interlingual" and "intralingual" instead. Here once again it turns out that delimitation is not always easy. This is particularly true in the case of orthography and phonology, but also holds good in the case of the other linguistic levels. Intrastructural interference is mostly attributable to wrong analogies and can be explained as being due to the fact that the linguistic data are not yet firmly enough established in the target language, thus giving rise to mistakes of overusage and too little usage of certain elements. Here it can be seen that E. Levenston takes too simplified and too one-sided a view of the teaching situation in assuming that related structures are in principle more frequently used than non-related ones[23]. No doubt the order in which structures are practised is also decisive. Otherwise quite a number of errors with regard to word order as well as use of aspect and tense, to mention just a few fields, may occur. In this connection the teacher is faced with the didactic problem of whether it is advisable to continue the practising of "basic patterns" up to the moment where one can be absolutely sure that even the worst pupil has grasped them, or whether there is not a risk of overdoing "drilling" in the case of the better and average pupils, which might then be a certain handicap in the learning of modified structures[24]. But this point can only be decided by way of experiment.

Moreover it should be necessary to investigate whether the tendency to overcorrectness and overgeneralization is not dependent upon certain types. Apart from the fact that the teacher's personality no doubt fairly often makes itself felt in this respect, we cannot exclude the assumption that particularly conscientious pupils are, in an excess of zeal, inclined to generalize and make excessive use of certain forms, especially when certain stylistic features of English are recommended to them with too much insistence, such as unfortunately happens only too often in the case of the English participial construction with "sentence linkage function". Owing to lack of information on certain variables, especially didactic ones such as knowledge of the pupil, of the teacher, the teaching material, the motivation, the objectives, and other conditions, we are not able to carry out a more refined analysis. Thus the scope of the present analysis is limited from the very beginning. We do, however, intend to improve it in the future, for further insights can only be expected from a close cooperation between research centres and schools and from the findings of other subjects such as pedagogics and sociology.

(*b*) *Evaluative lapsology:*

Here again we are concerned with an extremely difficult pedagogical and linguistic problem and have to rely entirely on the cooperation of our colleagues in the schools. From the previous section it should have become clear that contrastive linguistics should have some influence on error evaluation, though other aspects must be taken into consideration, too.

(*c*) *On error therapy:*

Making use of the knowledge of the main error sources, textbook authors and teachers will have to draw up adequate exercises. Phenomena such as "hypercorrectness" or "overgeneralization" will not be made use of without considerable forethought because some pupils regard these as representing linguistic reality and are, then, liable to make mistakes. But the chronological factor and the problem of age groups have also to be taken into account as well as the question of whether it is advisable to make the learner aware of differences and structures. We are convinced that in the case of adults the procedure of making them increasingly aware of the difficulties involved can definitely form part of the therapy. To what extent senior classes should be offered introductions to linguistics together with structural comparisons in the course of language analysis lessons as part of error therapy must be investigated experimentally. In connection with successful error therapy the setting up of groups homogeneous from the point of view of attainment would have to be taken into consideration, for, then, it would be possible to devise more uniform exercises for a particular attainment group. Curriculum research and research into measurement of attainment could make a valuable contribution to this sector.

Since many mistakes are attributable to a mixing of different stylistic levels — here the problems caused by the overlapping of British and American English must be included — the therapy will have to include stylistic exercises which make clear to the pupil the homogeneity of closed stylistic levels. Stylistic mistakes consisting in a mixing of stylistic levels are, of course, very difficult to explain in the foreign language, for the learner lacks the firm hold of the subject which would enable him to select almost instinctively the right level. Perhaps one can clarify the discrepancy more effectively by referring to analogous confusions between stylistic levels in the mother tongue than by offering unilingual exercises. Then the pupil would have to be familiarized via appropriate exercises with the linguistic elements of the various stylistic levels, or rather their location on a particular occasion. Moreover, in the course of such exercises it will have to be pointed out that the spoken and written language are based on different situations and consequently have to be handled in different ways. In view of the increasing significance of oral material in foreign language instruction it is striking that frequently, both with regard to language tests and error evaluation, none of the fundamental requirements necessary for assessing oral performance seem to exist and that for this reason there is often a disparity between the official goals on the one hand and the actual performance evaluation on an annual average on the other hand.

In view of the findings about the nature of an error described at the beginning, according to which an error is to be seen as a dynamic phenomenon and a process, and not as static and linear, it is an obvious conclusion that an automatic, repeated correction of parts or wholes is to little purpose. As a sensible alternative, language-teaching experts suggest that system-related repetition exercises be given as a revision of compositions written in class, where regularities are made conscious and practised by the learner through the means of paradigms of parallel forms[25]. Of course these exercises demand a great amount of the teacher's time and are difficult to carry out on account of the teacher's burden of work. Numerous teachers avoid the time-consuming procedure of prescribing written revision of composition, which also entails an enormous amount of correcting, and replace it by having revision done during subsequent lessons, in which case the formation of parallel forms can take into consideration the results of the previously written classroom composition. The procedure selected for exercises of this kind is a variant of "pattern practice". Then exercises are devised in the light of fairly superficial statistics of error distribution in the composition. Many experts are opposed to too narrow, non-contextualized corrective exercises of this kind, for corrective forms which are too isolated have the disadvantage that contextual conditions exceeding sentence units are not taken account of. It is self-evident that in the case of oral repetition of individual expressions so-called suprasegmental factors (stress, intonation, pauses, etc.) are not rendered correctly.

Thus in practising error therapy care must be taken that paradigmatic repetitions do not produce monotony, which would also place too much emphasis on particular units while others are neglected. The more independent and novel the way and the situation are by means of which the learner is made to arrive at the corrected linguistic form the more natural and unobtrusive will be the process by which he learns the form in question. Admittedly this procedure requires considerable empathy, linguistic knowledge, and much time.

We would be grateful if the discussion could be continued with the participation of teachers. The basic requirement for such a dialogue would, of course, be a reduction in the number of lessons to be given by the foreign-language teacher and appropriate in-service training facilities. In-service courses for teachers repeatedly demonstrate that questions concerning error analysis rank among the favourite subjects.

Notes

1. S. P. Corder, *The Significance of Learners' Errors, IRAL* 4 (1967), p. 161-171.
2. P. Strevens, *Two Ways of Looking at Error Analysis.* Paper presented at GAL Meeting (Stuttgart, November, 1969), p. 28-30.
3. V. J. Cook, *The Analogy between First and Second Language Learning, IRAL* 7 (1969), p. 207-216.
4. R. Brown and U. Bellugi, *Three Processes in the Child's Acquisition of Syntax.* In: *New Directions in the Study of Language*, ed. H. Lenneberg (Cambridge, Mass. 1964), p. 131-163.
5. Cf. V. J. Cook, *op. cit.*
6. S. P. Corder, *The Study of Error and the Problem of Correction.* Unpublished manuscript. To appear in: *Moderne Sprachwissenschaft und Fremdsprachenunterricht*, ed. G. Nickel.
7. Cf. V. J. Cook, *op. cit.*
8. C. C. Fries, *Teaching and Learning English as a Second Language* (Ann Arbor, 1945), p. 9.
9. R. Lado, *Linguistics Across Cultures* (Ann Arbor, 1957), p. 3.
10. J. C. Richards, *A Non-Contrastive Approach to Error Analysis.* Unpublished manuscript. To appear in: *English Language Teaching.*
11. G. Nickel, *Contrastive Linguistics and Foreign-Language Teaching.* To appear in: *Papers in Contrastive Linguistics*, ed. G. Nickel (Cambridge, 1971).
12. Cf. S. P. Corder, *The Significance . . .*, p. 162.
13. Cf. L. Newmark and D. A. Reibel, *Necessity and Sufficiency in Language Learning, IRAL* 6 (1968), p. 145-164.
14. On the history of this discipline see, e.g. G. Nickel and K. H. Wagner, *Contrastive Linguistics and Language Teaching, IRAL* 6 (1968), p. 233-255.
15. J. E. Alatis, ed. *Contrastive Linguistics and Its Pedagogical Implications.* In: *Report of the Nineteenth Annual Round Table Meeting on Linguistics and Language Studies/Monograph Series on Languages and Linguistics* No. 21 (Washington, 1968); G. Nickel, *Contrastive Linguistics and Foreign-Language Teaching.*
16. E. König and G. Nickel, *Transformationelle Restriktionen in der Verbalsyntax des Englischen und Deutschen.* In: *Probleme der kontrastiven Sprachwissenschaft. Jahrbuch des Instituts für Deutsche Sprache*, ed. H. Moser (Düsseldorf, 1970), p. 17-22.
17. Cf. R. P. Stockwell, J. D. Bowen, and J. W. Martin, *The Grammatical Structures of English and Spanish* (Chicago, 1965).
18. G. Nickel, *Rationalisierung des Fremdsprachenunterrichts, Paedagogica Europaea* 4 (1968), p. 92f.
19. *Ibid.*, p. 94.
20. Cf. G. Nickel, *Bericht über Ergebnisse der kontrastiven Analyse sprachlicher Phänomene im Deutschen und Englischen, Deutschunterricht für Ausländer*, Heft 5/6 (Munchen, 1970).
21. Cf. D. Dusková, *On Sources of Errors in Foreign Language Learning, IRAL* 7 (1969), p. 11-36.

22. Cf. G. Nickel, *Grundsätzliches zur Fehleranalyse und zum Problem der Fehlerbewertung*, *PAKS-Arbeitsbericht* 5 (Juli 1970), p. 2-30.

23. E. A. Levenston, *Over-Indulgence and Under-Representation — Aspects of Mother-Tongue Interference.* Paper read at the Second International Congress of Applied Linguistics. To appear in: *Papers in Contrastive Linguistics*, ed. G. Nickel (Cambridge, 1971).

24. In an essay forwarded to me by the Edinburgh linguist S. P. Corder (*The Study of Error and the Problem of Correction*) we read on p. 8: "It was found that whenever *is* occurred in the text the student was liable to add an *a* after it. Thus, *is a sitting, is a broken* were found beside *is sitting* and *is broken*. This learner was evidently operating a rule which said that whenever *is* occurs it may be optionally followed by *a*. We can only speculate that she developed this rule from excessive drilling of such sentences as *this is a book, this is a chair*, etc."

    Similarly it struck me that in learning German Norwegians often do not observe the rule of inversion in the subclause and use instead the normal word order of a main clause, though conditions in their mother tongue are similar to those in German. It is very unlikely that this is imported from the English language, as is often held, for the simple reason that some of the pupils who make this mistake have no knowledge of English. It looks as if the early and constant drilling of the basic pattern is responsible for this aberration.

25. Cf. F. Schubel, *Methodik des Englischunterrichts* (Frankfurt, 1966).

Discussion

Discussants: Nickel; Davies, Sibayan

The chief point of discussion was the explanatory capacity of contrastive linguistics. There have been claims that contrastive linguistics had some kind of total explanatory capacity; i.e. that all language errors could be accounted for by contrastive analysis, and that contrastive analysis could predict the degree of learning difficulty of given items. Experience has shown both these claim to be invalid. Many errors are unaccountable on the basis of contrastive analysis (e.g. errors arising from L_2 teaching or learning strategies), and "logical" predictions of language difficulties have been shown to be false. Nevertheless, contrastive analysis has a place as one among a number of explanatory aids for the analysis and prediction of learning errors.

Contrastive analysis tends to focus on errors, because for the theoretician, error is the best clue to what is going on in the language learning process.

LANGUAGE TESTS AND LANGUAGE ERRORS

by ALAN DAVIES

1. In this paper I first consider three questions: what is a test, what is an error, and what is error analysis. I then compare tests and error analysis and argue that while there are important theoretical differences it is not clear what information error analysis can provide which tests do not already give us.

2.1. What is a test? Numerous definitions have been given. Here are three:

1. 'An examination or test procedure, if it is to be efficient, must discriminate among the subjects examined or tested both reliably and validly.' (Pilliner, 1968, p. 23).
2. 'A psychological or educational test is a procedure designed to elicit certain behaviour from which one can make inferences about certain characteristics of an individual.' (Carroll, 1968, p. 46).
3. '. . . testing control of the (transfer learning) problems is testing control of the language.' (Lado, 1959, p. 24).

2.2. In the language field there are several kinds of test. They differ from one another in three ways, in purpose, in the group to be tested, and in the choice of language ("sample" might be a better term). The last two differences can be reduced to one; i.e. they depend on one another since we do not test someone who wishes to drive a car by putting him on a bicycle, nor do we test a class of French learners' knowledge of French by confronting them with a task in Chinese. A test, then, is a set of tasks (usually related ones since a test is structured round the principle of homogeneity) given to a group of learners. Here we have two sampling problems, the sample of learners and the sample of tasks. The first is difficult but can be settled to some extent statistically. Given that we know the population we are interested in, then we try, so far as possible, in making the test, to give it to all that population and in some cases we might be successful, e.g. all secondary school learners of Japanese in Australia — and even more easily if we restricted ourselves to, say, Sydney. But given that for most of the populations we are interested in we cannot get hold of everyone, then we must take a sample and it is accepted that if our sampling is random then we can make projections to what the results mean in terms of the population. The second sampling problem concerns the choice of tasks; if the first is difficult this is probably impossible. The problem here is not that of taking a random sample but of knowing what the population is in the first place. We do not know what are the discrete parts of language learning, nor do we know what their sequence is. What this means is that every test takes a set of bits of a language and assumes that they represent some part of what it is the native speaker knows (and/or does). It is helpful to distinguish four kinds of test, *viz.* Achievement, Proficiency, Aptitude, Diagnosis, and to do so in the following way, making use of the arrow of time: (Davies 1968).

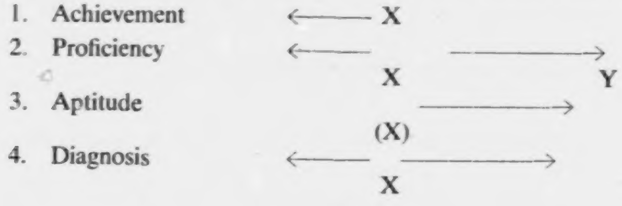


Figure 1

We have called these *kinds* of test; they may equally well be termed *uses* of tests so that it would presumably be possible to make use of one test instrument in all four possible ways. But this is straining the description somewhat; and for the moment we do not wish to argue about the similarity of content in these four types of test. What matters is that they should be seen as *uses* in the sense that the purpose for which they are being used is all-important.

2.3. In an achievement (or attainment) test we are concerned with assessing what has been learnt of a known syllabus. This may be within a school or within a total educational system: thus the typical external school examinations, the university degree examinations and so on are all examples of achievement tests. The use being made of the measure is to find out just how much has been learnt of what has been taught (i.e. of the syllabus). Achievement tests end there; however, though the primary interest may be in the past very often some further use is made of the same test in order to make 'meaningful decisions' about the pupils' future. It would, presumably, be possible to be interested entirely in the past of the pupils: Carroll's 'meaningful decisions' then would refer to the syllabus, i.e. to any necessary alterations to it that might be necessary or to the teaching method, etc., for the next group of students. But, as is obvious, achievement tests are almost always used for other purposes as well; it is important to recognise this and to account for it in one's test construction. But since an achievement test is a kind of shadow of the syllabus it is the latter that should have the predictive element built in. Unless the syllabus is designed so as to predict it is very difficult indeed for an achievement test to be made predictive.

2.4. Proficiency tests are concerned with assessing what has been learnt of a known or an unknown syllabus. Here we see the distinctions between proficiency and achievement. In the non-language field we might consider the Advanced Motorists' Test as a kind of proficiency test since there is the desire to apply a common standard to all who present themselves whatever their previous driving experience; over which, of course, there has been no control at all. In the language field there are several well known proficiency tests of the same journeyman kind: the Cambridge Proficiency examinations, the Michigan, TOEFL, and EPTB. These all imply that a common standard is being applied to all comers. More sophisticated proficiency tests (sophisticated in use not in design, of course) may be constructed for the all-important purpose of deciding just how much control over a language is needed for certain purposes (e. g. university studies in a second language). The proficiency tests just mentioned may well have started off as research measures to fulfil exactly such a purpose; it may be hoped that they did so for otherwise it is difficult to understand what type of control they are intended to represent.

2.5. If proficiency tests differ from achievement by the uncertainty as to previous instruction, it is in their relation to future needs or control purposes that they are distinguished from aptitude tests. Although this distinction is not always made in the literature it seems a useful one. We shall elaborate on the aptitude use in a moment; for the present let us conclude that proficiency in a language implies control over adequate skill in language for an extra-linguistic purpose.

2.6. Aptitude tests assess proficiency in language for language use. This is their distinction from proficiency tests. Hence, in Figure 1 the arrow related the language test under proficiency to *something else* (Y), and under aptitude to itself (X). An aptitude test is generally conceived as assessing amount of linguistic skill in itself for the purpose of learning languages. Generally it will mean for learning a particular language. It is to state the obvious to say that language aptitude is made up of a number of different skills, only one of which is (or may be) linguistic skill. Aptitude remains one of the most difficult areas in language testing research, mainly because of the reliance it must put on prediction. In Figure 1 we have represented the aptitude test X as (X) i.e. in brackets. We know where the arrow leads — to proficiency; and so aptitude research has to set up its goals first, must decide what is meant by particular language proficiency, i.e. the criterion. But what is not known is how this may be predicted, i.e. how success in learning, say, Russian, may be predicted before any Russian has been learnt. The attempt (and in a way the interest) in all aptitude work is not so much to predict future successful learners, though of course this has to be done, but to break down into component parts those skills in present linguistic control (often in the mother tongue) which relate to future control.

2.7. The diagnosis (more usually diagnostic) test belongs to a different category from the other three because it relates entirely to the use of the information and not at all to the presence of a skill in the learner. Achievement, proficiency and aptitude are all concerned with both use and skill. A diagnostic test is a use made by a teacher etc. of the information provided by one of these skills. A diagnostic test may be constructed for itself or it may be an additional use made of an achievement or proficiency test. If it is specially constructed it could perhaps be argued that some element of skill (or rather its absence) in the learner is being looked for. Indeed non-achievement may well be a sensible name for a diagnostic test. Just as the achievement test is built on the content of the previous learning, i.e. a sampling of the syllabus that has been taught, so the diagnostic test is built out of what has not been learnt. Like the proficiency test the diagnostic one looks both before and after: before to find out what is wrong with the previous learning, after in order to do something about it, i.e. for purposes of planning etc. remedial work.

2.8. To summarise in terms of this present paper for proficiency and diagnostic tests: a proficiency test makes some sort of 'best guess' of what learners should be able to do at some particular point. A diagnostic test is made up of 'errors' and here is the obvious link with error analysis. As we have seen they both could be *uses* of the same instrument. However, it is more usual to think of a diagnostic test as not being structured as a proficiency test since a diagnostic test can be made up of a number of unrelated items which it has been shown learners of a particular type tend to get wrong. It is also probable that the statistical rationale behind diagnostic and proficiency tests differs.

3.1. What does a test test? A test is not a questionnaire (though a questionnaire may eventually be the basis for a test) nor is it a survey; even less is it a set of exercises, or in any sense a teaching device. A test is a means of finding out something. The 'something' will depend on the use to which we may wish to put the test; we have already mentioned the main uses. The 'means of finding out' also requires some explanation. Like any 'scientific' test or experimental instrument, it must, as Pilliner tells us, be reliable and valid. Reliable means that any tester will obtain the same results; and there is the further implication that the testees will produce the same results on successive occasions. The 'objective' item is essentially an answer to the demand of reliability. A reliable test may be very efficient but not test anything very worth while. This is why a test must be valid too. Valid means that you know what you are testing and that it is what you want to test; this can be assessed by correlation or by content analysis techniques. There is the further implication of knowing what you want to test, namely that the ability under test conforms in some way to the theory of differential abilities, i.e. that it can be described on a normal curve. Along with reliability and validity we shall consider this as a third characteristic of tests.

3.2. Let us now consider some examples of test items which may suggest to us what it is a test tests:

1. *Reading*

   Underline the word closest in meaning to the underlined word:

   Our *tom* cat has been missing for a week.

   wild
   drum
   name
   male

2. *Listening*

   Here are 3 words. Are they all the same, all different or 2 the same?

   (bit, bit, beat)
   1 2 3

   All same
   2 same 12, 23, 13
   All different

3. *Listening*

   Hear are 3 sentences. Are all the same, all different or 2 the same?

   (It's 5 o'clock – read as statement, question, and with double stress on 5)

   All same
   2 same 12, 23, 13
   All different

4. *Listening*

   Listen to the following question and then choose the best answer:

   What did John leave with you?

   1. In the afternoon
   2. With his brother
   3. A large parcel
   4. Yes, he did.

5. *Reading*

Which of the following choices would a native speaker make?

my hairs
I'm going to wash the hair tonight.
my hair
hair

3.3. When we ask why such items as these are tested there are three answers:

1. they are critical points, either in Lado's sense of their being points of interference for the learner or that they are caused by analogy. Whatever the cause they have been shown up as Common Errors (French 1955).
2. they are easily testable. This is very important, perhaps the most important point. Producing the right utterance in the right situation is very important to the learner but useless to the tester because it is virtually untestable.
3. they are important to the learner at his particular stage, e.g. a physics lecture to advanced students in higher education. This is not a 'critical point' in general but is very important to a particular group. On the whole items will be determined on all three grounds. It is clear also that test items (point 2 above) test reception and production.

4.1. What is an error? In testing, errors are what a diagnostic test tests. But then there is the further question of how these are established in the first place and this may be on the basis of a survey of the 'Common Errors' type or simply on the basis of a previous proficiency test. What is the difference between items which are 'error'-full in a diagnostic test and the ordinary run of items in a proficiency test? Clearly we are not going to put items in a diagnostic test which are impossible for everyone in our target population. Surely there is no difference except that the diagnostic test items are more difficult for the average learner. Now just as we do not fill a diagnostic test with impossible items, so we do not fill a proficiency test with trivially easy items such that all learners in our target population get everything right. This may be good for our egos as teachers but would sin against our third characteristic of what makes a test; such a performance would do, perhaps, for an exercise, but not for a test. We conclude, therefore, that what a proficiency test does is just what a diagnostic test does, it produces errors. There will be, of course, a differential as to how many errors any one individual learner produces but this means that we have a test and not an exercise. It is also possible that in a diagnostic test we are more closely interested in a very homogeneous group and wish to distinguish individual errors and syndromes of error. In a diagnostic test our interest is psycholinguistic. We are now in a position to modify our definition of error. An error is what average and below average learners regard as a right response to a test item.

4.2. It follows from what we said earlier about the demand of reliability that the errors produced will be consistent ones. A test is, after all, not a trick. It does not capitalise on slips of the pen or tongue (or their correlates in reception). The errors produced will therefore be systematic ones in the sense that they will tell us something about the competence of the learners. We can, therefore, make use of a set of test results to make profiles of what a group of learners possesses as their competence, what is, in other words, their grammar and we can do the same for an individual learner.

5.1. We now come to our third question of what is error analysis. To an error analyst an error represents a deviance from native speaker competence. Corder (1969) of course points out that it is preferable not to speak of errors and to consider the whole production of a learner as potentially deviant, or, in his preferred term, idiosyncratic. I see his argument but it seems to me more reasonable to argue that if a learner produces a set of utterances which we must on all grounds, including contextual and situational, regard as non-deviant, we must ask ourselves how we judge them to be non-deviant and we must surely answer that it is in terms of their relationship to a grammar of native speaker competence. We consider errors therefore to be systematic deviances from native speaker competence and I should like to distinguish two senses of systematic: the first has to do with regularity; the second has to do with the system produced. The suggestion here, notably by Corder (1967), is that every learning stage provides us with a system: this implies an infinite number of possible systems. I see this as an attractive hypothesis but it does seem to me potentially vacuous since I can see no way in which we could ever describe an infinite number of systems or set up any kind of rules which would project them.

6.1. We are now in a position to make several distinctions between error analysis and tests. *First,* error analysis is individually oriented, it is more interested in psychological determinants than in group ones (though presumably error analysis could be done for a group with assumed common characteristics); tests are more concerned with group parameters and are therefore socially oriented. A social group is made up of individuals and individuals are of interest (unless they are poets) in terms of their language only if their behaviour can be generalised to some extent. We may state as the first difference between tests and error analysis that tests set up some norm of group language behaviour and then compare an individual's performance (or that of a group) with that of the norm. Error analysis in its investigation of individual language behaviour has no norm except assumed systematicness. But we must surely agree that it does also have the norm of native speaker proficiency in the target language with which it must implicitly compare the learner's behaviour.

6.2. *Second,* error analysis is a longitudinal method of study whereas testing is a 'sample bite' method. Of course, testing can be used longitudinally but tests still remain bites; this is also true of error analysis since no observation method is ever going to observe everything: that said, however, the more observed the better the sample on the whole.

6.3. *Third,* error analysis is open-ended (and probably data oriented); it asks: 'what do you say?'. Testing is objective (and probably non-data oriented); it asks: 'do you say X?' However, we must distinguish between orientation and technique since error analysis could easily switch from 'what do you say' to 'do you say X?' and would probably wish to do so by confining the learner to certain fixed responses; at this point it would become very much like tests in preparation; notice that I *am not* saying that it would become a test because I wish to distinguish clearly between tests in preparation when they are necessarily speculative and tests in use (i.e. already constructed ones) which have now become 'scientific' instruments.

6.4 *Fourth,* error analysis is research oriented; it is not committed to a 'normal', 'average' way of learning because it is concerned with the individual but this presumably means that there will eventually be useful generalisations, i.e. that there are patterns of learning. Testing, on the other hand, assumes knowledge of what has to be learnt and also that there is a statable progression norm of learning. Testing, therefore, makes a 'best guess' as to learning stages. A good language proficiency test could be used diagnostically to show patterns of learning, and could even be used sequentially (i.e. at successive stages in the learning process).

6.5 *In conclusion*, What have we been saying: error analysis is clinical; tests are *experimental*; error analysis is theoretical; tests are *practical*; error analysis is psycholinguistic; tests are *educational*. While testing is experimental it is a procedure for experiment, an instrument, whereas error analysis is in itself a whole view. While testing is clear about its methodology it is very uncertain how error analysis goes about its business. Error analysis and testing meet in two places: when error analysis is used as a technique itself (i.e. it has become 'do you say X?') and then is on the verge of a test; and when testing is itself a psycholinguistic operation.

6.6. In a word, then, error analysis is concerned with grammatical description; testing with placing individuals or groups within a grammar that has already been described. From this point of view it is testing that is data oriented and error analysis that is rule based. We may still wish to ask what rules error analysis can construct since they are potentially different for every learner. However, we have accepted that error analysis must be interested in eventual generalisation and therefore if error analysis could be made to produce statements as to learners' transitional competences in a useful way then the results would obviously be of great value to help in constructing more delicate or sophisticated, or simply just plain true Tests.

6.7. The basic difference between error analysis and testing therefore is that they make different underlying assumptions. Error analysis assumes that we do not know the grammar a learner has at any one time and proceeds to look for it (whether or not this can be done is another matter); testing assumes that we know what grammar the learner is aiming at (as does error analysis of course) and, much more important, that it is possible to suggest a limited finite number of ways that will take him there. We have seen that error analysis and testing agree about what has to be known (eventually); they disagree about the sequencing of arriving there. They further disagree as to how what has to be known, i.e. the language, may be broken down, or so it may appear. Error analysis will define its units in terms of what the system imposes. While a test will test, say, SV concord, or the negative transformation because they are there, error analysis will only incorporate them if they are relevant. However, this difference may be illusory because error analysis must start off by having the expectation that these features may be important; and testing may reject them as part of a test for a particular group of learners (one nationality, say) for whom these features may not discriminate or be sufficiently accounted for by other parts of testing.

6.8. Can testing be used to provide us with grammars for transitional competences? Testing could be used to show sequential learning and give sufficient number (infinite) of transitional competences. But the only change that a test can show is *growth*. Once reliability has been established further wrong responses to items previously got right must be treated as slips. (Incidentally, establishing reliability through item analysis is similar to discarding slips of the tongue, pen etc. in error analysis). Testing cannot handle, as error analysis claims to (by contrastive analysis), changes which may include shrinkage. Consider a test of 60 items on the English Verb, each with 4 choices. Clearly the purpose of this test (as of error analysis) is to establish the learner's grammar (actually, more likely to establish its distance from a native speaker's). Now, a test of only 9 items could produce over a million possible profiles or patterns of scores. It seems to me, therefore, unlikely that we shall ever arrive at either plausible grammars which are generalisable (60 items would produce enough individual combinations of different patterns for everyone in the world, present and past) or at acceptable ways of reconstructing our tests on each occasion to fit an individual's likely hypothesis.

6.9. But this is a pessimistic view. Testing cannot be as subtle as error analysis because it must be receptive while error analysis can be productive at least initially; I am not sure how long in technique it can remain productive. What this vast number of possible paths suggests surely is that testing, given good items with plenty of overlap, can be used for exactly the same job as error analysis. Granted that testing does make use of present assumptions as to grammar and sequencing; at the same time, since it sets up millions of conditions for any one state (progressively increased if we take states as a further multiplier), then testing surely provides the range of possible grammars that error analysis claims. We shall never, of course, analyse a million of these for tests but nor shall we for error analysis. What is pseudo procedure for one must be pseudo procedure for the other.

7.1. Let us push a little further this notion of what grammar testing uses and in particular what is the nature of the distractors that are used in testing. The purpose of Item Analysis, as we have said, is, in general, to make a test homogeneous; it maximises group patterns and avoids idiosyncratic ones; though there is dispute about this guillotine use of Item Analysis. Let us assume a test of, say, 20 items on Tense/ Aspect, all of which discriminate highly: here are 4 of them:

loves
1. John loved music whenever he hears it.
is loving

has lived
2. Now that summer has come, John is living here for a whole year.
has been living

Do you see
3. Did you see that bright star a moment ago?
Are you seeing

He had been studying
4. He has been studying in France for 3 years when war broke out.
He was studying

Let us consider an individual who responds half right, half wrong to all 20 items; what possible profiles can we suggest for him?

Here are three possible ones:

1. the choices under test (perfect, non-perfect, continuous, non-continuous are in free variation in his grammar).
2. perfect and non-perfect have coalesced as one or other.
3. continuous and non-continuous have coalesced as one or other, and so on.

7.2. Testing is inevitably limited by the range of distractors that it selects. Since it works to a known grammar and an assumed sequence it must use a limited range; what is more, its purpose of maximising group differences forces it to restrict and group the type of distractor it uses. Test results, therefore, can show only that the learner has a partial, not a different grasp, on the assumption that test items all make use of the grammar of the target language. But if they do more, by including distractors on the analogy of the learner's mother tongue or other languages he knows, or other possible hypotheses such as teacher's eccentricity, then we might be able to show by testing that the learner has a different and not just a partial grasp. The problem for testing is what hypotheses, and here is where we look to error analysis to provide some for us. But it remains difficult to see what can be the learner's hypotheses and how generalisable these may ever be if any of them are so idiosyncratic as to be concerned with, for example, teacher's eccentricity. Interference, yes; analogy, yes; other known languages, yes; and a good test will incorporate all these. We conclude, therefore that a test will give us information of *partial* knowledge not of *different* knowledge, i.e. that the profiles we arrive at will be those of free variation or of a gap in the system. We query whether error analysis, because of the sheer practical difficulty, can provide this information about *different* knowledge which it seeks to do. Our hope is, of course, that error analysis may come up with hypotheses which can be incorporated in new tests.

References

Carroll J. B. "The Psychology of Language Testing" (pp. 46–69) in *Language Testing Symposium*, ed. Alan Davies, OUP, 1968.

Corder S. P.

1. "The Significance of Learners' Errors", IRAL, 5/4, 1967.
2. Unpublished paper read at Dept. of Applied Linguistics, University of of Edinburgh, 1969.

Davies A. "Introduction" (pp. 1–18) in *Language Testing Symposium*, ed. Alan Davies, OUP, 1968.

French F. G. *Common Errors in English*, Longmans, 1955.

Lado R. *Language Testing*, Longmans, 1959.

Pilliner A. "Subjective and Objective Testing" (pp. 19–35) in *Language Testing Symposium*, ed. Alan Davies, OUP, 1968.

Discussion

Discussants: Davies; Cherrier, Nickel, Otter, Palmer, Saha

Discussion of error analysis brings up the question of success analysis. The focus of testing tends to be on error, yet it is only diagnostic testing that should so focus; achievement testing should focus on success. There is here a tactical problem, however. Error shows where success is *not*, and hence by implication where it *is*. The value of error is that it can be localized; success is too vast. Success analysis obviously offers a field of research that needs to be explored, but the way into it seems to be through error analysis.

One requirement for fruitful research in this area is the development of language learning models. Language teachers have listened too much to theoretical linguists, who have tended to set up linguistic models possibly because applied linguists have never done so. But the relevance of linguistic models to language learning is at times tenuous. Various language models may offer ways to get at a language learning model, and a pedagogical grammar can be a mixture of various models (unlike a strictly linguistic model). The transformational-generative model may have something to offer, but in the field of phonology it is clear that while generative phonology is useful in linguistic theory it is of no use to the language teacher, to whom taxonomic phonology *is* useful.

In the development of a language learning model, there is a special role which language teachers might play. Good language teachers (i.e. those that get effective results) base much of their teaching on their intutions of what should be done in given situations, rather than on theory. They may often pay lip-service to a linguistic or methodological theory which they do not in practice actually follow. Their practice, however, is demonstrably effective. These practices of good teachers ought to be taken into account as a sort of intuitional, heuristic basis for the construction of a science of language pedagogy and a language learning model.

ERROR ANALYSIS AS AN INDEX OF CLASS PROGRESS

by LIM KIAT BOEY

Many kinds of standardized English language tests are in existence both in America and Britain but these have been standardized for the students in those countries, whose mother tongue is English. There are also standardized tests of English as a second language, but again they are mostly for native speakers of European languages. (In fact, this highlights one of the limitations of tests based on Lado's theory of language testing. His basic assumption is that a second language test must be based on a contrastive analysis of the first and second language.) Standardized tests of English as a second language specifically meant for the countries of South East Asia are still few and far between. I hope one of the outcomes of this seminar will be the beginning of efforts to build such tests. Meantime language teachers will continue to devise their own tests, as they have always done. Even if standardized tests were available, however, in the teacher's day-to-day work, he needs to give frequent small diagnostic and attainment tests in order to get feedback and evaluate the effectiveness of his teaching. Such tests are perhaps best devised by the teacher himself, as I will show in what follows.

One of the problems in language teaching is to get an objective measure of how far a class has really assimilated a certain usage or structure without resort to a standardized test. Such a measure would also give an indication of the degree of effectiveness of the methods used with the class and provide the clues to modifications and developments. This paper is a record of an attempt to find such a yardstick by means of error analysis.

At its inception, the then School of Education was involved in an English teaching project at Sekolah Alam Shah, a Malay-medium secondary school. The aim was to raise the level of attainment in English in the Sixth Form, so as to enable the students to compete on equal terms with students from English medium schools. This was the first Malay-medium Sixth Form and it consisted of 35 students, all Malays except for 2 Chinese boys. There were 8 girls in the class. Two members from the School of Education started work with the class towards the end of September, 1963.

First of all, a series of tests was devised and given to the class in order to find their level of attainment in English. Among the tests given was one on written comprehension. A passage was taken from "Precis Practice for Overseas Students" by Bright and Nicholson, suitable for Form III in an English-medium school. However, the questions were modified and two other sections were added. Section A required full statement answers in order to get the students to produce sentences. Section B consisted of objective comprehension of the True/False type and Section C a vocabulary test of the multiple-choice type. All the questions in the three sections were based on the same passage. The nature of the test meant that only the answers in Section A would yield errors in usage and therefore only this section was used in the analysis. The impression that the marker of the papers received was that the areas of greatest weakness were:

1. Use of passive voice.
2. Omission of copula.
3. Tenses.
4. Use of articles.
5. Concord.
6. Prepositions.

Working on this impression as a basis, when actual teaching began in October, the two members from the School of Education devised various drills and exercises to eliminate these common errors, as well as to develop linguistic skills in other directions. It was found that following drill in one area, e.g. articles, the students would show marked improvement in that area when tested specifically in that area. But in their free composition or when they had moved on to a different kind of exercise, one noticed reversions to wrong usage. This is not very surprising, since language habits die hard and language is very much a matter of habit. However, it was felt that it would be helpful to have an objective picture of the progress of the class and to find out if the methods used were producing any results. One way of finding out was to give another written comprehension test at the end of the term, parallel to the one given at the beginning, and compare the students' performance in the two tests.

Accordingly, another passage written by the same author was taken from "Precis Practice for Oversea Students", of a standard comparable to the first passage, and questions set on similar lines. When the papers were marked and the raw scores compared, it was found that every student except one had improved his score. Was this to be taken as an indication that the class had made some progress in English? Or was it due to other factors, such as an easier passage or easier questions or loss of initial nervousness? If one assumed they had improved, in what directions had they done so? Was it in comprehension or production?

To get a clearer picture, it was decided to make a detailed analysis of the mistakes made in production. Table I (p. 110) shows the results of the first test and Table II (p. 111) the second. The classification of mistakes explains itself in the headings, except for "verb", which means the wrong part of a verb has been attached to an auxiliary e.g. "have destroy", "did not climbed", and "by climbed". Some mistakes may come under two or three different headings, for example, "have pass" may be classified under "verb" and again under "concord" if the subject is singular. A few other mistakes have not been entered in the tables as they are isolated ones.

Several interesting things came to light. First of all, on the impressionistic view, it was thought that mistakes with regard to the passive voice and copula came higher in the list than the article and preposition, but the order in Table I is as follows:

1. Spelling.
2. Tense.
3. Preposition.
4. Article.
5. Passive Voice.
6. Vocabulary.
7. Plural.
8. Sentence Construction.
9. Copula and Concord.
10. Verb.

Secondly, the order for the number of students who made the mistakes does not correspond with the order for the number of mistakes.

Thirdly, comparing the third and fourth rows from the bottom of Table I with the corresponding rows in Table II, one finds a decrease in the number of errors and number of students making the errors except in three areas — verbs, vocabulary and sentence constructions, where there is an increase. Can one therefore conclude that the students have improved in the areas in which they had been drilled? Such a conclusion cannot be based on these figures on the errors alone. It must be in relation to the amount written, for the more one writes the more opportunities there are to make mistakes. So a count was made of the number of sentences written and the number of sentences that are error free. The last two columns of each table give the figures, which show a difference of 4 sentences in the total number written. To compare the 2 sets of figures from Table I and Table II, they are converted to a percentage of the number of sentences written as shown in Table III.

	Spelling	Plurals	Passive Voice	Verb	Tenses	Concord	Omission of Copula	Vocabulary	Preposition	Articles	Sentence Construction	Error free sentences	Total No. of sentences
Table I	19.3	8.2	11.1	5.7	17.3	6.7	6.7	10.1	16.4	15.4	7.2	28.9	207
Table II	5.4	2.9	2.9	7.3	11.3	1.9	5.4	16.2	5.4	3.9	9.3	40.3	203

Table III Comparison of Errors in the Two Tests

The figures in Table III show a decrease in mistakes made but an increase of errors in three areas. What is the significance of this? The greatest increase is in vocabulary. This is not surprising for though part of the teaching time was devoted to building up their vocabulary, the test was not based on the vocabulary learned. The increases under sentence construction and verb are small. Sentence constructions, like vocabulary, as they occur in answers to comprehension tests, tend to be influenced by the passage given and the questions set, and may or may not show the effects of drill. There was no drill on verbs except in relation to tense. But neither was there specific drill in concord except in connection with drill in the use of the present tense, nor drill on the plural and the copula except in connection with drill on the article, and yet there is a decrease in errors in the areas of concord, copula and the plural. The biggest decrease in errors is in the area of spelling, but spelling performance is again related to the words that occur in the passage. If 2 or 3 words commonly misspelt occur in the passage, then the total number of errors increases. The big drop in mistakes in the use of prepositions can only be accounted for by incidental correction, for there was no specific drill on this. There are three areas in which there was explicit drill and which show a corresponding decrease in wrong usage. These are tenses, passive voice and articles. In these three areas the improvement is probably due to the drill and exercises, because both passages are written in the past tense, both offer opportunities to use passive voice, and articles, of course, are always in evidence.

What then, are the conclusions one may draw from such an error analysis? How far can it be used as an index of class progress? It is clear from the above analysis that the following precautions must be taken:

1. The passages chosen must be of comparable difficulty and the questions must be parallel.

2. The number of sentences analysed for errors should be equal, and if not, the figures should be reduced to a common basis before comparison can be made.
3. Not all items lend themselves to this purely quantitative analysis. Certain items such as vocabulary and spelling are more dependent on the content of the passage. Such items may need a further analysis, where the variety as well as spread of the errors in that category are considered, in order to get a fair picture of the situation.

If such caution is exercised, then such an analysis will give the teacher a picture of class progress, albeit the index is a negative one. And this brings us to the uses of such an analysis:

1. It shows the areas in which the class is weak.
2. Using the analysis as a basis, the teacher can decide whether the class as a whole will benefit from drill in a certain direction, or whether there are certain marked groups which require drilling in different aspects and divide the class accordingly. This will work towards the greatest economy in the use of time and energy for the teacher and class and also obviate the danger of boredom.
3. A detailed analysis of those items which call for such a study will reveal the necessary steps to be taken by the teacher. For example, an examination of the spelling errors may show that the errors are of the type which result from uncertainty as to when to double the consonant e.g. "writter", "dinning". The teacher could then give the class the rule with regard to this or devise a drill to deal with recalcitrant words.
4. Where the figures show a general improvement, the analysis may act as a sort of reinforcement to the teacher, who is often discouraged by the seeming imperviousness of the class, when errors keep recurring. Or it may rouse the teacher to examine himself and his methods, where the figures show a complete lack of response.

It will have been noticed that the tests of comprehension we have used may be found in many textbooks. It will also have been noticed that the classification of errors presented here is a gross one. The aim was to demonstrate a method of testing, its problems and its uses to the majority of teachers, who have had little or no training in modern linguistics. But the analysis can be as detailed and sophisticated as the teacher has the talent, time and inclination for. Under sentence construction, for example, one could make a finer analysis and classify the errors into: dangling participles, misuse of certain sub-ordinating conjunctions, incomplete sentences, run-on sentences, etc. If the teacher is familiar with the techniques of transformational grammar he may choose to classify the errors of sentence construction into those attributable to surface structure and those to deep structure. The exercises devised to remedy the situation will then be based on insights provided by transformational analysis. In fact, the more knowledge a teacher has of the structures of the first and second language (irrespective of whether he has learned traditional grammar, structural grammar or transformational grammar), the better will he understand the sources of errors through his analysis and how to help his students overcome them.

Although the index of class progress suggested here is a negative one and although it takes time, it has its uses, and such an analysis from time to time may be salutary for the teacher.

Discussion

Discussants: Lim; Carroll, Davies, Durr, Mohan, Nababan, Ng, Nickel, Saha

In error analysis, whether carried on formally as a special technique or informally as part of the daily conduct of classes, the teacher who is competent in the students' native language as well as in English has an advantage in perception of the sources of errors and possibly greater sympathy for the students' problem. Contrastive analysis, however, can be helpful even to the bi-lingual teacher.

Any error analysis faces the fundamental problem of defining "error." Error has been defined (by Prof. Davies) as "any deviation from native speaker competence." But native speaker competence is often an unattainable ideal, and we should consider the possibility of defining our standard of correctness. In India, it has been decided to set up an "Indian standard English"; "error" would then be defined as a deviation from this standard. This is not to say that some kind of substandard English is here being accepted. "Indian English" is proposed as a standard English meeting the criterion of intelligibility internationally as well as within India. It is of course not easy to describe accurately this Indian standard English, but it is certainly possible. The same approach might be taken in Singapore, or the Philippines.

There are some objections to this approach, however. The standard of intelligibility must be strictly maintained, and this may be difficult especially in the near future when progressively fewer native speaking teachers will be available to maintain standards. From the point of view specifically of testing, one might still maintain the concept of error as a deviation from native speaker competence, but modify the expected standard to meet the requirements of what is effective within the student's situation.

For the analysis of error, any basic theory of grammar may be used; the teacher should use whatever kind of grammar he is familiar with, and he can carry out the analysis to any degree of precision and sophistication he has the energy, inclination and time for. The errors in the testing program under discussion here were analyzed in terms of discrete language components. It should be noted, however, that the test was a "global" test, not a test of specific items, because it had been found that students, after drilling, often responded accurately in tests of discrete items but reverted to previous errors in the context of free response.

The scoring of a free production test of this type may offer some problems. Research in progress under the direction of Dr. John Carroll indicates that the best way of expressing the score seems to be in terms of the proportion of the number of correct clauses to the total number of clauses produced.

TABLE I

Number of Students	Spelling	Plurals	Passive Voice	Verb	Tense	Concord	Omission of Copula	Vocabulary	Preposition	Article	Sentence Construction	Correct Sentences	Total No. of Sentences
1.	1	—	—	—	2	—	—	—	1	1	1	3	6
2.	—	—	—	—	—	—	1	—	—	2	—	—	2
3.	—	—	—	1	3	1	1	1	1	1	1	1	5
4.	3	—	—	—	2	—	—	—	1	2	—	3	7
5.	—	2	—	—	—	—	3	—	2	1	1	1	7
6.	4	—	2	—	1	1	—	—	—	—	—	1	6
7.	1	1	3	—	4	—	2	1	1	2	1	2	10
8.	—	—	—	—	2	—	—	—	—	3	1	—	4
9.	—	1	—	—	—	—	—	1	1	2	—	3	6
10.	1	1	1	1	—	1	—	—	1	—	—	2	5
11.	3	1	1	—	1	—	—	1	1	—	—	2	6
12.	—	—	2	1	2	2	2	2	1	—	1	—	5
13.	2	1	—	—	1	—	—	—	1	1	—	2	5
14.	2	—	1	1	1	—	—	—	—	—	—	1	4
15.	—	—	—	—	2	—	—	—	2	2	—	3	7
16.	—	—	—	—	—	—	—	—	1	1	—	4	6
17.	3	1	2	—	—	1	—	1	2	1	—	1	7
18.	—	—	1	—	—	—	—	—	—	1	—	5	7
19.	1	2	—	—	1	2	—	—	—	—	—	2	6
20.	1	—	—	3	—	—	—	—	2	2	—	—	4
21.	—	1	1	—	1	1	—	—	2	—	1	—	6
22.	3	—	2	1	—	1	—	2	1	4	—	—	5
23.	—	—	—	—	—	—	—	—	—	—	—	5	6
24.	1	—	1	—	1	1	1	2	—	—	—	2	7
25.	1	—	—	—	—	—	—	—	2	—	1	2	7
26.	—	2	—	—	—	1	—	2	1	—	—	3	6
27.	—	1	1	—	—	—	—	—	2	—	1	1	6
28.	—	1	—	1	2	1	—	1	1	1	1	1	6
29.	1	1	—	1	2	—	—	2	3	1	—	1	7
30.	3	—	1	—	—	—	1	1	2	1	—	2	7
31.	2	1	1	1	3	—	2	2	1	—	1	—	6
32.	1	—	—	—	2	—	—	—	—	1	—	3	6
33.	3	—	—	—	1	—	—	—	—	1	2	2	6
34.	—	—	—	—	1	1	—	2	—	—	1	1	5
35.	3	—	3	1	1	—	1	—	1	1	1	1	6
Total No. of Errors	40	17	23	12	36	14	14	21	34	32	15		
Total No. of Students with Errors	20	14	15	10	21	12	9	14	24	21	14		
Total correct sentences												60	
Grand Total													207

TABLE II

Nunber of Students	Spelling	Plurals	Passive Voice	Verb	Tense	Concord	Omission of Copula	Vocabulary	Preposition	Article	Sentence Construction	Correct Sentences	Total No. of Sentences
1.	—	—	—	—	—	1	1	2	—	—	—	4	7
2.	1	—	—	1	3	1	—	—	1	—	1	1	5
3.	—	1	—	—	3	—	—	—	—	—	1	—	5
4.	—	—	—	1	—	—	—	1	—	—	1	3	5
5.	—	—	—	—	2	—	1	—	—	1	1	—	7
6.	—	—	—	—	—	1	—	1	—	1	—	2	5
7.	1	—	—	—	3	—	—	2	1	—	—	2	7
8.	3	—	—	—	—	—	1	—	—	2	1	3	10
9.	1	—	—	2	2	—	1	1	—	—	1	—	6
10.	—	1	—	—	—	—	—	—	—	—	—	4	5
11.	—	—	—	—	1	—	—	5	1	—	—	—	6
12.	—	—	—	1	—	—	—	1	1	—	2	2	5
13.	—	1	—	1	1	—	1	3	1	—	1	—	5
14.	—	—	—	1	—	—	—	1	—	—	1	3	6
15.	—	—	—	1	—	—	—	2	2	—	1	3	7
16.	—	—	—	—	—	—	—	—	—	1	—	3	5
17.	—	—	—	1	1	—	—	—	1	1	2	2	5
18.	—	—	1	—	—	—	—	—	—	—	—	5	8
19.	—	—	—	—	—	—	1	—	—	—	—	3	6
20.	—	—	1	—	—	—	—	—	—	—	1	4	5
21.	1	1	3	—	2	1	1	—	—	1	—	—	6
22.	—	—	—	1	1	—	1	2	—	—	—	2	7
23.	—	—	—	—	—	—	—	1	—	—	—	4	5
24.	—	1	—	2	—	—	—	2	2	—	—	—	5
25.	—	—	—	—	—	—	—	—	—	—	—	5	5
26.	—	—	—	2	1	—	—	1	—	1	—	—	5
27.	—	—	—	—	—	—	—	—	—	—	—	5	5
28.	—	—	—	—	—	—	—	1	—	—	—	4	5
29.	1	—	—	—	2	—	—	2	1	—	—	1	6
30.	1	1	—	—	—	—	—	—	—	—	1	1	5
31.	—	—	—	—	—	—	1	2	—	—	1	3	5
32.	—	—	—	1	—	—	—	—	—	—	—	6	7
33.	—	—	—	—	1	—	—	—	—	—	1	4	6
34.	1	—	—	—	—	—	1	1	—	—	1	2	6
35.	1	—	1	—	—	—	1	2	—	—	1	1	5
Total No. of Errors	11	6	6	15	23	4	11	33	11	8	19		
Total No. of Students with Errors	9	6	4	12	12	4	11	19	9	7	17		
Total correct sentences												82	
Grand Total													203

ENGLISH LANGUAGE TESTING FOR UNIVERSITY ENTRANCE

by ROBERT B. KAPLAN

In an essay entitled "The Retreat from the Word," George Steiner writes:

> The Apostle tells us that in the beginning was the Word. He gives us no assurance as to the end.
>
> It is appropriate that he should have used the Greek language to express the Hellenistic concept of the *Logos,* for it is to the fact of its Greek-Judaic inheritance that the Western civilization owes its essentially verbal character. We take this character for granted . . . We live inside the act of discourse. (But) there are modes of intellectual and sensuous reality founded not on language, but on other communicative energies such as the icon or the musical note. And there are actions of the spirit rooted in silence. It is difficult to *speak* of these, for how should speech justly convey the shape and vitality of silence? But I can cite examples of what I mean.
>
> In certain Oriental metaphysics, in Buddhism and Taoism, the soul is envisioned as ascending from the gross impediments of the material, through domains of insight that can be rendered by lofty and precise language, toward ever deepening silence. The highest, purest reach of the contemplative act is that which has learned to leave language behind . . . It is only by breaking through the walls of language that visionary observance can enter the world of total and immediate understanding. Where such understanding is attained, the truth need no longer suffer the impurities and fragmentation that speech necessarily entails. It need not conform to the naive logic and linear conception of time implicit in (English) syntax. In ultimated truth, past, present, and future are simultaneously comprised. It is the temporal structure of language that keeps them artificially distinct . . .
>
> Pascal is nearer the mainstream of classic Western feeling when he says that the silence of cosmic space strikes terror . . .
>
> The primacy of the word, of that which can be spoken and communicated in discourse, is characteristic of the Greek and Judaic genius and carried over into Christianity. The classic . . . sense of the world strives to order reality within the governance of language. Literature, philosophy, theology, law, the arts of history, are endeavors to enclose within the bounds of rational discourse the sum of human experience, its recorded past, its present condition and future expectations . . . All truth and realness . . . can be housed inside the walls of language.[1]

As repositories and transmitters of that tradition, universities in the United States are themselves engulfed in verbalism, and as a result of that fact demand that those who enter to pursue a truth housed inside the walls of language be themselves housed within those walls.

In other words, colleges and universities within the United States believe that all students who enter must have an adequate control of language, specifically of English, as a condition of admission. In itself, this fact is derived from reasonable assumptions; that is, if indeed knowledge (and, one supposes, truth) is transmitted verbally, then it is reasonable to require the learner to be attuned to the language in which it is transmitted in order to receive it, or at least in order to profit from it in a purely pragmatic sense. However, as is the case in a great many other philosophical issues, there is a vast gap between the reasonable hypothesis and the logistic accomplishment.

---

[1] Reprinted in *Language and Silence: Essays on Language, Literature, and the Inhuman* (New York: Atheneum, 1967), pp. 13, 14, 15.

Leaving the philosophical, and turning toward the practical, may I call attention to a number of significant points which appear to have been ignored. The greatest effort among colleges and universities in the United States has been directed toward the so-called "developing" countries and even more specifically toward the middle levels of civil service and of the private sector in those countries. But this audience contains in large part exactly those individuals who have had neither the leisure nor the financial capability in many global areas to develop relatively high-level proficiency in the English language. Furthermore, in many global areas, English is taught as a *foreign* language, as a purely academic subject, rather than as the medium of communication of a living people, so that even if the specified audience had had the leisure and the money to acquire training, the training may not have been available. Thus, the requirement of English language proficiency as a condition of admission tends to eliminate exactly the audience which most colleges and universities in the United States wish to attract.

Additionally, even assuming that language proficiency were a valid screening device, there are other problems. First, what is meant by language proficiency? What kinds of skills really are involved in demonstrating competence in a language? The quadrivium of speaking, listening, reading, and writing does appear to be measurable to some extent, but does the measurement of these four skills have any significant relationship to language proficiency? There are whole vast areas of communication untouched by these four "linguistic" skills. If one thinks in terms of Guilford's model of intelligence, it becomes apparent that the measurement of these four skills, in relation to an index of language proficiency, merely scratches the surface. Indeed, in an unpublished paper entitled "Taxonomy of Testing Objectives" (November 23, 1967), E. Glynn Lewis proposes a preliminary model for which he isolates five broad categories of linguistic concern, four of psychological concern, and three of social concern. To these he adds another set of two geographic concerns; thus he posits a pair of models each consisting of five by four by three cells as his first stage. Then he isolates five categories of testing objectives and calls attention to the problems inherent in maturation; thus his second stage model (in which he groups students into five generalized age categories) consists of five by two by five cells. In order to get a reasonably accurate generalization of proficiency, then, each of the cells in both models would need to be filled and the resultants of the two sets would need to be correlated. The state of the art hasn't arrived at anything like this.

There is no question that a number of extant measuring instruments provide some sort of index to proficiency in the English language. But again, there are those unfortunate logistic problems. Probably the best available proficiency test is the Test of English as a Foreign Language (TOEFL). This test is periodically administered throughout the world. Logistically, there is frequently a lag of anywhere from twelve to eighteen months between the time any given student takes the test at any given point of administration and the time that same warm body arrives on a college or university campus in the United States. Even if the index thus provided is accurate, there is every likelihood that significant change in either direction may occur in that time interval. Suppose, for example, that a candidate for admission to a school in the United States travels from a provincial university to the capital city, studies English intensively for three months, takes the TOEFL, and then returns to his home to await reply; and suppose that in his home city where he works as an official in a municipal office, there are no native speakers of English and only a handful of speakers of English as a second language. Suppose further that this mythical student waits one year for his notification of admission. It is fairly clear that the proficiency of the individual at the time that he arrives on the campus of the school that admitted him may be vastly different. This example ignores the difficulty he may have experienced in getting to the center at which the examination is administered, the difficulty he may have undergone in finding the funds for the three months of intensive study, and so on. It also ignores the possibility that he may for whatever reason simply have found a more fluent friend to take the examination for him, or contrarily that he may himself retake the test any number of times at relatively brief intervals thereby introducing a "practice effect" which will affect the test score but not the real proficiency, whatever that is.

Quite aside from these issues is the issue of the nature of the test — indeed of a great many proficiency tests; that is, such tests, as a function of grading simplification and standardization — even as a function of cost — attempt to measure only passive skills, only listening and reading. Success in a college or university in the United States may depend, at least in certain academic disciplines, more upon the ability to speak and write. But even if it were possible to overcome the complexities in scoring samples of connected discourse both written and oral, the other manifestations of the problem — the time lag, the human complications, and so on — would still exist.

The fact that the measuring instrument is a proficiency gauge in itself constitutes quite another difficulty. Ignoring for a moment the other issues raised, it is clear that a proficiency test measures the amount of information or skill an individual possesses at a given moment in time. It is not intended to provide information showing how long and at what cost the individual has worked to arrive at that point; nor is it designed to show how soon or at what cost the individual may acquire any one additional linguistic item. A proficiency test does not, and is not intended to, measure aptitude, but some indication of aptitude might allow an admitting institution to reach an intelligent decision concerning the length and the extent of additional language training a given individual might need to study in a school in the United States.

Since I speak as a test consumer rather than as a test producer or as a psycholinguist, I must confess my nearly total ignorance in regard to a number of related variables. However, there is no doubt in my mind that attitude and motivation are extremely significant factors in the whole question of language acquisition. While I am generally aware of some of the work being done in this area by scholars like Wallace Lambert, I confess that I am not aware of the state of the art nor of the possibility of developing broad-scale instruments to do something with these variables.

From another point of view, there are at least two additional problems. The first of these relates to the fact that, to the best of my knowledge, no meaningful correlation has been established between achievement on a language proficiency test or in an intensive English course on the one hand and academic success in a college or university on the other. Institutions in the United States operate on the assumption that high-level English language proficiency is the *sine qua non* of academic success. Of course, "academic success" is an extremely difficult concept to measure. Identifying it with grades both constitutes a circular argument and begs the question. It is circular in the sense that grades are assumed to be themselves an index of success; it begs the question in that it ignores important differences in the meaning of a given grade at graduate vs undergraduate level and among disciplines as disparate as physics, literature, and economics (all of which may be simultaneously present in the study of an undergraduate student in the United States.) Even ignoring the difficulty of defining 'academic success," there is no clear evidence either that there is or that there is not a correlation between second-language achievement and academic success.

The second of these problems is related to a political issue; that which has come to be called, in common parlance, "The Braindrain," and officially, "the international migration of intellectual talent." The issue is simply this: relatively large numbers of individuals who come to the United States to study tend, for a variety of reasons, to remain there and thus to frustrate the expectations of those individuals and agencies who sent them. The factors which cause individuals to remain in the United States are many and varied. There is no question that greater economic possibilities, or the promise of them, play a role. But there is also no question that language instruction and the accompanying acculturation are also important factors. It is clear that the more successful an individual is in increasing his competence, the more likely he is to become deeply acculturated. Consequently, the more deeply acculturated he becomes, the more difficult the task of returning to his own culture appears, and the greater is the probability that he will rationalize other reasons to allow him to remain in a situation in which he has come to be comfortable. This tendency is not inherently "bad"; in fact, there are those who maintain that it is "good." Its goodness or badness is irrelevant. The point is that the result contributes to additional political tension in a world which is already possessed of more tension than it knows what do with.

In summary then, the assumptions of English language testing for University entrance in the United States are predicated upon a number of misconceptions; namely, that it is possible to determine rather finitely what given language competence is, that it is desirable to do so, and that it is necessary to do so. Quite the contrary, as I have tried to show, it is not desirable to do so because doing so will screen out exactly the population which colleges and universities in the United States hope to serve. It is not presently possible to do so because proficiency tests do not provide finite information concerning language competence, and instruments which may provide such information are not presently available for this population.

It is not even necessary to do so because a relatively large number of colleges and universities in the United States have the capability to provide intensive language instruction to those who seem to need it. Granting all the reservations already stated about the relationship between second-language competence and academic success, there does seem to be some empirical evidence that students who study the English language intensively after arrival in the United States do perform somewhat better in academic work than those who do not. To what extent this phenomenon is a manifestation of TLC (tender-loving-care), or of the provision, through an intensive program, of a kind of linguistico-cultural decompression chamber, it is difficult to say. An argument can, I think, be made for *second-* as opposed to *foreign-*language instruction. (I am using Ferguson's distinction, in which second-language instruction occurs in the milieu in which the target language is native, while foreign-language instruction occurs in the milieu in which the target language is foreign and thus the instruction is academic.) Whatever the reasons, the intensive course on arrival seems to be more efficient. Please note that these comments are in no sense intended to be critical of English teaching around the world. My comments are related to rather specific terminal objectives for a rather specific audience, under rather special circumstances, being trained in what may correctly be called a specialized restricted code. It may be that these factors — the objectives, the register, and the audience — are not always clear even among those who are most deeply concerned in second-language instruction, and that as a result semantic confusion sometimes replaces intelligent curricular planning; thus, in the extreme, the teaching about English literature may be assumed as equivalent to intensive language instruction. But, I wander. My point is that instruction on the admitting campus in the United States may be more efficient than other kinds of instruction. That is not to say that we do it better; that is to say that we do it under better circumstances. Obviously, however, such a recommendation again raises all sorts of logistic problems. It increases the cost of education. It lengthens the student's stay in the United States. It increases the probability of greater acculturation and of alienation from the home culture. It increases the flow of wealth out of the home country and into the host country. It increases the strain on family ties and on various loyalties. But for all that, the amount of increase in all these areas may not be that great, and I for one tend to believe that the increased efficiency tends at least to balance out the increased costs and dangers.

In summary, then, it is not necessary to establish high-level English language proficiency for university entrance in the United States because institutions in the United States have the capability to provide instruction which appears to be more efficient in the long run.

While I am reasonably certain that everything I have said so far is as accurate as I can make it, the pragmatic fact is that my argument is largely academic. Most colleges and universities in the United States do indeed require high-level English language proficiency as a criterion of admission. And most use one or another of the available proficiency tests to determine the admissibility of a student. In fact, some place greater emphasis on this particular criterion than they do on academic ability or achievement as demonstrated by a variety of other kinds of measures. As you are undoubtedly aware, there are presently three principle tests in wide use — the "Michigan" Test, the ALIGU — American Language Institute, Georgetown University, and the TOEFL. In addition to these, Jack Upshur, when he chaired a conference on language testing at the University of Michigan in 1967, pulled together a list of 150 tests; Buros' *Sixth Mental Measurement Yearbook* lists 23 tests (no attempt was made to survey earlier editions of Buros), and Jean-Guy Savard's recent volume entitled *Bibliographie Analytique de Tests de Langue* (Universite Laval, 1969) reviews 57 tests. Obviously, these tests vary in validity and in reliability, not to mention quality. New tests appear periodically. Just before leaving my office last week, I received an announcement of the publication by McGraw-Hill of a new test by David P. Harris and Leslie A. Palmer called CELT — A Comprehensive English Language Test. A kind of cycle seems to have developed: schools require language proficiency tests; extant tests do not seem to be adequate; there is a growing market for tests; more publishers get into the act; more schools require language proficiency tests, and so on *ad nauseum*. There is neither the time today, nor, I gather, the interest, to review this vast array of instruments. Suffice it to say that, though I consider it somewhat immoral, schools in the United States are likely to continue to require high-level English language proficiency as a criterion of admission for the foreseeable future. Schools in the United States are likely to continue to use proficiency instruments even though these instruments quickly lose their security, are administered at times and in places which may be inconvenient to candidates, and may measure the wrong skills.

Given these conditions, I am inclined to offer a bit of advice to prospective candidates. Since, unless a candidate applies to only one school — a heavy hazard — he is likely to be required to take several different tests, I would recommend studying English, rather than studying "for the test." Since the candidate is unlikely to apply only to one institution, I would recommend including among the applications at least one to an institution offering an intensive English program. Since there is some evidence that English language study on arrival is to some degree more efficient, I would recommend that every candidate simply plan to spend a semester or a year, over and above the time required for any given degree program, in language acquisition, and that both candidates and sponsors accept the increases in cost.

About two years ago, my colleague, Robert A. Jones, and I invested considerable time and effort in trying to factor out the variables operating in a Cloze Procedure Proficiency instrument in English as a Second Language. The only variable we could clearly identify in that situation was sex. In reporting our results, we indicated that we were prone to recommend a horizontal study of certain prominent features of that variable. In the end, this may be the only constructive idea I have to offer.

Discussion

Discussants: Kaplan, Moehnilabib; Carroll, Davies, Lee, Nickel, Otter

It was observed that many of the tests used for university entrance to American universities contained a strong emphasis, even an over-emphasis, on vocabulary. Admittedly, however, many of the language difficulties of students involved vocabulary, which therefore needed to be tested along with other elements.

Apparently, also, language tests are interpreted as measures of intelligence in some way, as well as specifically of language proficiency. These tests are being interpreted not by linguists but by admissions officers for the purpose of making decisions on admission, and intelligence data is being looked for. There is a correlation of approximately .7 between language proficiency tests and general intelligence tests, although it is not clear whether this is due to a basic correlation between language and general intelligence or merely due to the fact that language plays so large a role in intelligence testing. The correlation is closer also for L_1 proficiency then for L_2. At any rate, admissions officers can hardly be faulted for looking for intelligence in incoming students.

Various kinds of language tests are used for admissions purposes. Some test general language proficiency, some include tests of special registers. Specialized faculties (e.g. science) often do not recognize the need for basic general language ability but insist strongly on proficiency in the specialized register. Yet experience in the United States and United Kingdom tends to indicate that a basic minimum of general language proficiency is required for success in any field of study but that beyond that minimum the student's success seems to depend on his ability in his field of specialization rather than on special language competence.

Although it was suggested by Dr. Kaplan that proficiency tests for students coming to English-speaking universities be generally eliminated, it was pointed out that as long as students come to the United States or United Kingdom for stays of no more than six or nine months, this would not be feasible. It seems, however, that in future fewer students will come to the United States for short-term courses, and more for longer courses.

STANDARDIZED ENGLISH TESTS FOR UNIVERSITY ENTRANCE

by M. MOEHNILABIB

Introduction

The present paper discusses some problems in connection with the development and use of standardized English tests for University entrance in Indonesia. Standardized testing has not been used widely in the teaching of English in Indonesian schools and universities. In spite of many serious efforts to improve English teaching materials and methods, and to improve the quality of English teachers at all levels, the method of testing has received very little attention. In general, the importance of a good testing method for the success of teaching seems not to have been adequately understood by most English teachers in High Schools and Universities. This is evident from the low quality of most classroom English tests and more general kinds of examinations such as High School final examinations and University entrance examinations. These tests can be so highly subjective that two independent scorers may give to the same paper scores as different as "good" and "poor". No evaluation of the validity and reliability of the test items has ever been made, and the scores have never been given an objective interpretation in terms of some acceptable norms or criteria.

High School final examinations, of which English is a part, are considered very formal. The Ministry of Education and the local High School inspectorates assume full responsibility in the planning and construction of the examination questions as well as in the administration of the examinations. For the administration local Boards of Examiners are set up annually, of which the members are usually teachers from various High Schools. However, in spite of the highly formal nature of the examinations and the organizational complexity in their construction and administration, one of the most important aspects of an examination, namely the reliability and validity of the test items, has been completely ignored. No serious attempts to improve the quality of High School examinations have been made until this year. Currently a team of English teachers at one of the Indonesian Institutes of Education has been assigned to plan and develop a battery of standardized achievement tests of English for High School students, which in the next few years is expected to replace the traditional forms of High School final examinations.

Unlike High School final examinations, University entrance tests, of which English is a part, are the full responsibility of each University or Institute, both in their construction and in their administration. The general situation of University English entrance tests is slightly better than that of High School final examinations. In at least 14 State Institutes of Education in various parts of Indonesia, objective questions have been used whenever appropriate, although no full-scale standardized tests have been successfully developed and administered. Thorough familiarity with principles and techniques is still lacking among many instructors and test makers. In many cases the objective type of question used in the English entrance tests possesses a low degree of validity and reliability. The great amount of work and expenditure needed for the construction of good standardized tests has also been an obstacle. In spite of all the problems, however, attempts have been made at one of the Indonesian Institutes of Education to develop and use a valid set of standardized English entrance tests. The discussion in this paper will focus on the problems involved in these attempts.

Standardized English Entrance Tests

The idea of developing and using standardized English entrance tests in some Indonesian Universities was partly due to the standardized English tests developed by Robert Lado at the English Language Institute of the University of Michigan, which early in the 1960s were made familiar to some Indonesian Universities. Objective English entrance tests began to be developed on the model of Lado's Tests. However, the development of fully standardized tests has met all the problems and obstacles mentioned earlier, so that up to now the results have still been tentative and a lot of revisions must be made.

Construction of Tests

The first problem in the planning and construction of the standardized test concerns various factors that must be considered in determining the aspects of English to be tested, the level of the test, and the suitable types and forms of the test questions. One of these factors is the learning problems Indonesian students are expected to encounter in studying English at the university level. Simple contrastive analysis between English and Indonesian structures cannot be expected to show all the learning problems, because to most Indonesian students Indonesian is the second language they acquire at school and the influence of their mother tongue (one of the many local languages in Indonesia) is very strong on their command of the Indonesian language. This situation has greatly complicated the problem of deciding what must and must not be included in the test. In addition to that, a consideration must also be given to the students' level of English mastery as the result of 6 years of English instruction at the Junior and Senior High Schools. This cannot be based simply on an inspection of the High School English curriculum, because in the last ten years the curriculum has undergone various changes and the implementation of the curriculum also varies greatly from one school to another. A nationwide survey of the implementation of the High School curriculum is needed to provide a sound basis for judging the students' level of English mastery.

Another is the problem of emphasis on the testing of discrete elements of English and the testing of integrated skills. It is not always easy to determine the degree of emphasis that should be given to each in order to obtain a balanced and objective picture of the students' overall mastery of English. A certain amount of arbitrariness is inevitable. Thus the amount of time for each section of the test and the number of items in each section are not always determined by linguistic and other relevant factors. Various other factors such as economy of administration and scoring, balance in the physical appearance of the test as a whole, etc. may have some influence.

Economy of administration and scoring is of great importance in the planning and construction of the test although it has very little to do with linguistic and pedagogical principles. A test, however good linguistically and pedagogically, can be of no use unless the factor of economy is taken into account.

The purpose of the English entrance tests is another factor that affects the planning and construction of the tests. The English entrance tests in the English Department have the main purpose of selection and admission. Thus on the basis of the grades in the tests a candidate may or not be admitted to the English

Department. The importance of this purpose is evident from the fact that in the English Department of most Indonesian Universities only about one out of eight candidates is admitted. The second purpose of the English entrance tests in the English Department is diagnostic. Thus the results of the tests serve as guide-lines for minor changes and adjustments in the English curriculum. On the other hand, the English entrance tests in the other Departments are designed for those who will need English only in relation to their major fields of study. The only purpose of English entrance tests for these students is diagnostic. These different purposes of the tests must be taken into consideration in the planning and construction of the tests.

On the basis of the factors discussed above the coverage of the tests is determined. This includes the areas of English to be tested, the balance between the testing of discrete elements of English and the testing of integrated skills, the number of test questions and the amount of testing time. The last two points, however, are also determined through the try-out of the tests. The entire test (excluding the Production part of Pronunciation, and Speaking) contains 330 test items that must be finished in less than 200 minutes. These are divided into the following sections:

I. Discrete elements of English:		
1. Pronunciation: (Recognition)		
(*a*) Vowels and Consonants	30 items	12 minutes
(*b*) Stress (Word and Sentence)	20 items	8 minutes
(*c*) Intonation	10 items	4 minutes
2. Structure: (Recognition and Production)		
(*a*) Morphology (Inflection)	20 items	8 minutes
(*b*) Syntax	80 items	32 minutes
3. Vocabulary: (Recognition and Production)	70 items	30 minutes
II. Integrated skills:		
1. Reading Comprehension	40 items	30 minutes
2. Auditory Comprehension	40 items	30 minutes
3. Writing	20 items	30 minutes
4. Speaking (and Pronunciation)	—	(10 minutes)
Total number of test items:	330 items	194 minutes

All the tests of discrete elements include both recognition and production tests. In the Pronunciation test, however, only the recognition part is given above. For economy of administration the production part is incorporated into the Speaking test. The Pronunciation test includes a representative number of test items on vowels, consonants, word stresses and some sentence stresses, and some items on intonation patterns. The test on Structure also covers both recognition and production. It includes test items on morphology (mainly inflection) and syntax, which includes tenses, pronouns, articles and other function words, word order, and structural patterns.

In the testing of integrated skills the order of the skills presented here is somewhat different from the one commonly accepted in foreign language teaching. This is simply a reflection of the order of emphasis on the four language skills recognized in Indonesia. Reading Comprehension and Auditory Comprehension present no specific problem in administration as well as in scoring. But Writing and Speaking pose problems in scoring and in administration. For this reason the testing of these two skills has not been given proportionally important places in the standardized English entrance tests.

From the number of test items, it seems that testing of discrete elements is given a far stronger emphasis than that of integrated skills. But from the amount of testing time for each section, the testing of integrated skills receives a much stronger stress than that of discrete elements. As has been mentioned earlier, the decision as to which of the two should receive a stronger emphasis is inevitably arbitrary. However, if ways could be found to better control the accuracy and objectivity of the testing of integrated skills, it is reasonable that this should receive a greater emphasis.

In order to achieve high accuracy and objectivity as a measuring instrument, the types of test questions in the standardized English entrance tests must be given serious considerations. A fault or weakness in the types or forms of test questions could reduce or even eliminate the value of the test items in question. Therefore suitability of each test question for each type of problem is imperative. On the other hand, economy of administration and scoring to a certain extent also determines the choice of the types of test questions. For the recognition tests of various language elements and skills, some objective types of question, notably the multiple-choice type, are suitable. But one must not conclude that objective test questions are automatically valid. The validity and reliability of each test item are to be evaluated individually and carefully.

The actual preparations of the standardized English entrance tests discussed here involve the following steps:

1. Survey of the High School English curriculum and its implementation in order to determine the general level of the students' mastery of English. Unfortunately the survey could only be conducted on a very limited scale and the result must be considered very tentative. It is used in conjunction with some other data of the High School graduates' achievement in English.
2. The listing of all learning problems of Indonesian students studying English on the basis of some contrastive analysis between English and Indonesian structures. Some of the difficulties in this step have been mentioned earlier; namely, that Indonesian is not the first language of most Indonesian students.
3. The writing out of the test questions, keeping in mind the suitable types of questions and answers for the standardized entrance test.
4. Tryouts of the test: First on a group of thirty Fifth-Year students of the English Department, as the nearest available substitute for native speakers. This tryout eliminates a number of ambiguous and confusing test items. The second tryout is conducted on about 80 First-Year and Second-Year students of the English Department, and on about 250 High School graduates who are candidates in the English Department. The tryout is repeated again on the candidates admitted to the English Department four weeks after the first tryout in order to determine the reliability of the test. The third tryout is conducted after an item

analysis of the test results has been made and invalid test items discarded. Using the new First-Year and the Second-Year students (about 60 in number) this tryout has the aim of finding the suitable number of test questions and the right amount of testing time.

5. The final editing of the standardized English entrance tests is made after the tryouts are finished. The tests consist of the following physical parts:
   (*a*) Test booklet, containing separate sections of Pronunciation Test, Structure Test, Vocabulary Test, Reading Comprehension Test, Auditory Comprehension Test, Writing Test, and Speaking Test (the last is to be given orally by a qualified teacher).
   (*b*) Directions for administering the tests.
   (*c*) Answer sheets.
   (*d*) Scoring keys.

Use of Tests

The standardized tests were first used for the entrance examination in the English Department of the Institute of Education in Malang, Indonesia, in 1968 with about 240 candidates, and in 1969 were used for 204 candidates of the English Department of the same Institute. Comparison has been made between the result of the tests and the students' later achievement as shown in the promotion examination at the end of the first year. In the 1968 entrance tests 55 out of the 240 candidates were accepted. Of the 55 First-Year students, 48 were promoted at the end of the school-year. Detailed comparison of the result of the entrance tests and the students' grades in the promotion exams shows a high correlation. No comparison has been made between the result of the 1969 entrance tests, which will be made at the end of the second semester in 1970.

It is hoped that the comparisons will show the success or failure of the standardized English tests developed.

PREDICTING INDIVIDUAL DIFFERENCES IN ENGLISH LANGUAGE LEARNING

by JOHN B. CARROLL

No doubt there are individual differences in most kinds of learning situations. For years, psychologists have shown that some cluster of abilities called "intelligence" or "scholastic aptitude" have important relations with overall success in school. From the thousands of studies that have been published on this topic, it appears that school success is to a considerable extent dependent on the student's mastery of his native language, on his ability to reason, and his ability to think in quantitative terms. In addition, it is dependent upon his motivation, or at least the amount of time he actually spends on his studies in relation to the amount of time he would need to spend in order to achieve at his maximum.

Foreign languages are becoming even more important in most school curricula. Two questions can be raised; Are individual differences in foreign language learning dependent upon the same abilities as other school subjects? Are students motivated to study foreign languages in the same way that they are motivated to study other school subjects? In general, the research that I have done suggests that the answers to these questions are in the negative. Foreign language learning is something special: success in learning depends upon a somewhat different set of abilities than other school subjects, and students are not motivated to study foreign languages in the same way that they are motivated to study other school subjects. In short, individual differences in foreign language learning are not the same as individual differences in the learning of most other school subjects .Therefore we will have to treat them separately. The case of foreign languages is somewhat similar to the case of music: everybody knows that musical ability is something rather special: musical talent does not automatically go along with high intelligence, and musical talent can indeed be found in some people who have only meager aptitudes for success in school in general. In the same way, talent for foreign languages does not automatically accompany high intelligence, and it may appear in some individuals who are not otherwise very successful in school. I am not suggesting, of course, that musical talent and foreign language talents are the same; in fact, I think that in general they are very different, contrary to some fairly widely held opinions. Persons who combine musical and foreign language talents do exist, but from a statistical point of view this is a coincidence.

The proposition that there are individual differences in foreign language learning is not exactly self-evident, but it is abundantly supported by the common experience of teachers. Let us limit ourselves to the case of the person who is learning a foreign language at some time *beyond* the time he learns his native language. The matter of individual differences in the learning of the native language has interesting but different problems which will not concern us here. But if we take groups of children, say in the third grade, or in the seventh grade, or at the high school level, and particularly if we take groups of college students or adults, we find wide individual differences in the success they have in learning a foreign language — even if they are equally motivated and are given the best kind of instruction we know how to give.

Conceptually, I find it desirable to think of these individual differences as differences in the *rate* at which the person can acquire the foreign language. It is not an accident that we often speak of some persons as *fast* learners and others as *slow* learners. For if we can allow each person to learn at his own rate, we will find that the rates of learning do vary widely. In fact, the evidence suggests that rates of learning are distributed like many other human traits, that is, according to so-called "normal", bell-shaped frequency distribution. That is, the majority of people learn at more or less average rates, while there are some who learn either much faster or much slower than the average. Up at the top of the distribution are a few people who may be called "geniuses" at learning foreign languages; these are the people we meet once in a while who seem to be able to acquire a foreign language almost overnight. At the bottom of the distribution are a few people who are virtual "idiots" as far as learning foreign languages is concerned. They may be brilliant in something else, like mathematics or poetry writing, but one may expect them never to get far in foreign language learning.

Several further observations may be made at this point. First, an individual's rate of foreign language learning is more or less constant: if he is a slow learner, he is slow at the beginning of foreign language study and also he is slow later on. It is rare that one observes a slow learner actually pick up speed, and if one does, it may be because of some extraneous factor, such as a suddenly increased amount of motivation and effort; in this case, the person was probably not a slow learner after all. Second, rates of learning, for a given individual, are approximately the same regardless of what language he is studying — aside from the inherent differences in the difficulty of languages that apply to all learners. That is, a person who is slow in learning French will be expected to be slow, relative to other learners, in learning any other language, whether it is Spanish (which is sometimes erroneously supposed to be an "easy" language for Americans) or Russian (which is usually somewhat harder for Americans to learn than some other languages). Third, rates of learning probably do not change much over the course of one's life. The evidence for this is slim, actually, but common observation suggests that it may be true. Fourth, we do not know how to change an individual's characteristic rate of foreign language learning. Little research has been done on this question, unfortunately, but I will have some suggestions about it later on.

These individual differences in rate of foreign language learning may be spoken of as differences in foreign language *aptitude*, but in talking about foreign language aptitude I hope you will not infer that it corresponds to some bump on the head or some way in which the brain is constructed that causes good or poor learning. Exactly what foreign language aptitude is, in physiological terms, is a mystery. I don't know whether any aspect of it is inherited, or what aspects of it are learned. There is a little evidence to suggest that certain aspects of it are indeed influenced by heredity, but this is very problematical. I would not want to leave the impression that foreign language aptitude is, on the whole, inherited. It may have some basis in very early learning on the part of the young child. In any case, we have to take the fact of individual differences in foreign language learning as a "given". There are three things we can do about these differences. The remainder of this talk will discuss them.

The first thing we can do about individual differences in FL learning is to try to measure and predict these differences before the person actually undertakes foreign language study. I have spent a good deal of time doing research to make this possible, and as you may know, I have been able to develop a practical measuring instrument called the *Modern Language Aptitude Test*. This instrument is designed for testing persons from the ninth grade up through adulthood, but it can also be used with the brighter students of the seventh and eighth grades. A version of the test now published is designed for elementary school children in the third to the sixth grades. The senior form of the test has been validated on thousands of cases, and is being widely used. I am not alone in the endeavor to make tests to predict rates of foreign language learning: Paul Pimsleur has just published a test called the *Pimsleur Language Aptitude Battery*, based on somewhat the same principles as mine but designed for children in the 7th to 12th grades. There are one or two other foreign language *Aptitude* tests either on the market or restricted to use in certain organizations like the U.S. Army. The tests are not all equal in validity, but the very fact that it has been possible to construct and validate them seems to prove that differences in foreign language aptitude can be measured and predicted.

Furthermore, the nature of the tests gives quite a few leads as to the nature of foreign language aptitude, at least its psychological components. In constructing my elementary school version of the MLAT, I was surprised to find that the kinds of tests that proved most valid and useful were very similar to the kinds that had proved most valid and useful at the upper level.

Let me emphasize at this point that I believe foreign language aptitude is complex — that it depends upon a number of rather separate and specialised traits of the language learner. Of course, *rate* is one entity — unless you try to analyze foreign language learning at an almost microscopic level of detail, on the whole the individual's progress is fast or slow according to the net sum of his various language aptitudes. Nevertheless, it is interesting to examine the separate capacities that make up foreign language aptitude as a whole and that go into the prediction of the rate at which the learner will progress. I have identified four main abilities of this sort.

The first of these is something that probably has no parallel or counterpart in intelligence testing, and I always find it difficult to explain, perhaps because I am still not quite sure what it is, even though I can measure it in a number of different ways It is something I have been calling *phonetic coding ability*, but before you draw too much out of that phrase let me try to explain. Basically, it seems to be the person's ability to apprehend a particular speech sound or combination of sounds and, at some later time, identify it, recognize it, or recall it as different from some other sound or combination of sounds. Phonetic coding ability is *phonetic* because so far as I know it applies only to stimuli that can be considered as speech sounds. It does not, for example, apply to the apprehension and recall of groups of spoken digits as in the usual memory span test, and certainly it does not seem to apply in the case of materials that can be presented and remembered in purely visual terms, like alphabetic letters. Phonetic coding ability has to do with *coding* because I imagine that the person who is good at this ability is somehow able to "tag" or "code" the speech stimulus in the process of storing it in his memory. If you don't want to accept this aspect of my characterization, perhaps you will be satisfied if I call it simply *phonetic memory ability*. You can see how important such an ability is in learning a foreign language. In the early stages, one has to learn to recognize and also pronounce a series of foreign sounds, and this can be done best by a person who is inherently good at recognizing and remembering particular speech sounds. At later stages, one has to recognize and remember whole groups of sounds; for the person who is good at phonetic coding, perhaps these sounds are apprehended as single impressions, i.e. as *Gestalten*. This ability can be measured in a number of ways. Perhaps one of the best ways is to use a test that has to be administered individually: we pronounce a nonsense syllable or two, or perhaps a short phrase in a foreign language, and then give the subject a little mental arithmetic to do for about ten seconds before he is asked to repeat the sounds that he has heard. The delay is inserted in this test to make it necessary for the subject to store in memory the sounds he has heard; he is not allowed to repeat them on a purely imitative basis. This test seems to work well in the few instances when I have tried it, but it is of no use if the requirement is for group testing. We have to contrive various dodges and subterfuges to make this ability show itself in a group test. One such dodge is to make the individual learn a new set of printed symbols for the sounds of his own language; I have tried having the subject learn a completely new alphabet, like the Devanagari alphabet used in Sanskrit and Hindi, but the best expedient seems to be to make him learn a new phonetic transcription using mainly the Roman alphabet. This sort of test is administered as a group test as a tape recording (the Phonetic Script test, Part II of the MLAT). Another dodge is to see whether the individual has acquired ready responses to phonetic-orthographic stimuli in his native language: in my Spelling Clues test, the individual has to recognize very rapidly the words represented by somewhat abbreviated, partially phonetic spellings of English words. There are still other ways of measuring phonetic memory, enough to suggest that it is an ability of wide yet subtle importance in dealing with language stimuli.

A second ability of major importance in learning foreign languages is what I have called *grammatical sensitivity*. Fundamentally, it is the ability to recognize — that is, be aware of — the grammatical functions of words and other grammatical elements in sentences, even in one's native language. Now to be sure, everybody who speaks English as a native language has somehow acquired, subconsciously, some kind of competence with the grammar of that language, in the sense that he automatically uses the syntax of his language in such a way as to create understandable sentences. But not everybody, it seems, can bring this automatic competence to the level of awareness. Even linguists have trouble writing the grammar of their own language, let alone other languages. Even though nearly everybody has been exposed to training in formal grammar at one or more stages of his school career, not everybody can perform certain linguistic tasks that depend upon his perceiving grammatical functions and relationships. Even training in formal grammar does not seem to "cut through" this inability. In one good test of grammatical sensitivity, used in my MLAT, the student has to find words or phrases that have similar grammatical functions in two different sentences. For example, if I give you two sentences: *John gave Mary an apple*, and *Tom's brother interviewed Mr. Smith last Friday*, the task is to find what word or words in the second sentence have a function that parallels that of *apple* in the first sentence. To do this, the subject does not have to know any terminology of formal grammar. He does, however, have to be able to make a conscious analysis of the grammar of the two sentences. Again, it is easy to see how such an ability would be relevant in foreign language learning, no matter what role grammar plays in the instruction, because the student does have to work out some sort of grammatical analysis of the foreign language he is learning.

A third major component of foreign language aptitude is another kind of memory ability — rote learning ability for the meanings of foreign language words and expressions. It is often noticed, in studies of paired-associate learning in experimental psychology, that students differ widely in their ability to acquire the meanings of a list of nonsense syllables in a short time; experimental psychologists find that back of this ability seems to lie some kind of facility for making use of the associations that one has for the things that are to be connected. Apparently this sort of ability comes into play in learning a foreign language. That is, one of the problems the learner has is that of "connecting up" the foreign words and phrases with meanings and concepts he has already acquired. This is true regardless of how a foreign language vocabulary is taught. Some foreign language teachers believe that vocabulary should be taught only "in context", and they advise students to avoid the use of vocabulary lists, flashcards, and the like. Perhaps this advice is wise — I am not persuaded that it is always wise, but even when vocabulary is learned in context, that is, in the course of learning dialogues or reading prose passages, there is still a problem of connecting the arbitrary foreign language sound patterns with meanings of some sort. We can rather easily test this ability by giving the student a short vocabulary list of foreign words and their meanings, then testing him on his retention after a very short time.

A fourth major component of foreign language ability is inductive language learning ability. It can be best measured by giving the subject a series of sentences in a foreign language (which could be an artificial one) constructed and sequenced in such a way that it is possible to work out their grammar. My colleague in the development of MLAT, Stanley Sapon, worked out such a test, given with film strip, tape recorder, and test booklet. In fact he attempted to model the test after the most highly-approved audio-lingual teaching procedures. Some subjects "picked up" the grammar of his artificial language "Tem-Tem" very readily, by noticing the changes in words that accompanied changes in grammatical meanings, all well illustrated in changes in the pictures that accompanied the sentences spoken on the tape. Others did not seem to understand how to work out the grammar of Tem-Tem; in fact, perhaps they did not understand how a language is put together. Although the scores on this test were highly predictive of success in learning a real language, such as Chinese, we could not include the test in the commercial battery because it took too much time and equipment to give it. At any rate, it did provide us with insight as to one of the components of foreign language aptitude.

We did not attempt to measure motivation for foreign language study, but this is an important variable in foreign language study. It affects mainly the amount of time that the individual is willing to spend in practising and learning. The critical variable, I say, is the time the individual is willing to spend in learning; it does not matter very much *how* he is motivated — whether to get good grades or to learn the language for some utilitarian purpose he has in mind. Pimsleur, in his language aptitude battery, asks the student to rate his interest in foreign language, and finds that this is indeed one of the valid predictors of success. Sapon and I did not include this variable in the test score because of the possibility that the individual would deceive either himself or the examiner as to his true motivations.

Sometimes language aptitude tests such as mine have been criticized because they do not present, it is claimed, good "models" of language learning procedures. For example, the rote memory type of vocabulary learning test is criticized because it is not typical of the way students are advised to learn foreign language vocabulary. In reply, I would first point out that an aptitude test is not necessarily designed, and does not have to be designed, as a model learning situation. The purpose of an aptitude test is to measure as reliably as possible, and in the shortest time possible, the abilities and traits that underlie success in learning a foreign language. Doing this has nothing to do with presenting good models of language learning procedures. In fact, it is even possible that a test that concentrated on providing good models of language learning would fail to capture and measure the abilities that are normally relevant in language learning.

Now, I said that there are three things one can do about foreign language aptitude. I have discussed ways of measuring and predicting it. The second thing one can do about it is to try to adapt instruction to differences in foreign language aptitude. There are various ways of doing this. The most obvious one is to select only individuals with rather high language aptitude and teach only *them;* or one could move the cutting point down on the scale and use the test mainly to screen out the students who have a high likelihood of failure. This is the way the test is sometimes used, either by schools, or by organizations such as the Air Force, the Peace Corps, or the Foreign Service, where failure in foreign language training is costly both to the individual and to the organization. Use of the test as a selection device depends, then, on the practical situation and also to some extent on one's educational objectives and philosophy. Some would say that since the test is not perfectly accurate in identifying those who will fail, everybody should be given a chance to try learning a foreign language. I will grant, as a matter of fact, that despite the high validity of the test in predicting rate of learning in a variety of situations, probably the best indicator of success in learning is a practical try-out of learning provided the student is well motivated and he really is given sufficiently good instruction over a long period to allow his strengths and weaknesses to show themselves.

If the test is not used as a selection device, it can often be used as a guidance instrument, that is, as a measure that will show the individual with reasonable accuracy what his chances are of making good in foreign language study. Then the decision as to whether to study a foreign language is up to the individual. If the test shows that he has poor chances of being successful, but he still wants to try, he should by all means be given the opportunity to try.

A related issue, by the way, has to do with the use of the test as a way of "excusing" a student from a language requirement. Although the test has been used in this way in some colleges, I do not recommend its use for this purpose except with certain qualifications. For one thing, it is admitted that the test is not *perfectly* valid; it occasionally makes mistakes, for one reason or another. Second, a student could easily malinger on this test, i.e. just not try as hard as he could. I would recommend, rather, that the test be used as a diagnostic instrument. It could be given to a student who is failing his foreign language study: if he does well on the test, he would certainly be required to stay with his FL requirement. If he does poorly on the test, in conjunction with failing in his FL courses, the decision as to whether he should be excused from FL study would have to be based on a sensitive clinical judgment about how well motivated he was in taking the test, how well motivated he is in FL study, and other factors.

The chief way in which a FL aptitude test could be used in normal situations, it seems to me, is as a predictor of the rate at which the individual could successfully master a foreign language. It would therefore be used as a means of setting up sections that would go at different rates, or it could be used as a means of individualizing instruction so that even within a given class, some students would be allowed to go ahead much faster than the average student, and other students would be allowed to progress more slowly than the average. Highly apt students can be given advanced tapes and workbook material to study by themselves, with occasional help from the teacher, and under certain circumstances they might be able to skip over a semester or a quarter and be placed in a more advanced section than normal. Slow students could be given extra help, or programmed instruction materials that would allow them to work very slowly; they might be allowed to take a year to cover the ground normally covered in a semester. The remaining students would constitute the majority of students and they would all progress at approximately the same rate. Such a system, or something like it, would prevent what so often occurs — namely, the situation where the progress of a class is determined by the learning rates of its slowest members. Likewise, it would allow the language learning "geniuses" to capitalize on their gift. Some of these could easily complete a three-year course in two years or even less. I have had reports from teachers that sectioning students by ability in foreign language, whether measured by an aptitude test or on the basis of past performance, makes language teaching much more successful for the students and pleasant for the teacher. Of course, in small schools where there are limited possibilities for sectioning this plan has some administrative complications. Nevertheless, I am convinced that the obstacles can be overcome.

A third way in which a foreign language aptitude test can be used has already been touched on — as a diagnostic instrument. I cannot claim that it has been deliberately designed for this purpose, and in fact, the diagnostic use of the test is somewhat limited by the fact that the subtests are somewhat short and limited in reliability. Nevertheless, I suggest that attempts be made to use it in this way. I have indicated some of the aspects of foreign language aptitude that we have tried to measure with MLAT; often the pattern of scores that an individual makes on this test will indicate where specific weaknesses lie. For example, a person who makes relatively low scores on those parts of the test measuring phonetic memory may need special help in learning to remember foreign sounds and their combination; a person who makes a relatively low score on the Words in Sentences test may need special help in learning foreign language grammar; and a person who does relatively poorly on the last test, the one of rote memory for vocabulary, may need special help in finding devices to help him remember foreign language vocabulary. It is even possible that special help given to an individual in the light of weaknesses shown on an aptitude test will, in the long run, improve his foreign language aptitude and thus accelerate his rate of progress. This is an area where research is much needed.

In conclusion, it seems to me that the area of foreign language training is a particularly promising one for taking account of individual differences in a meaningful way. Aptitude can be more easily and accurately measured than in some other areas of the curriculum, and the content of foreign language courses naturally grades itself in difficulty and complexity. I would not claim that adapting instruction to individual differences will make it possible for every student to be equally successful in the end, but I do think that it might at least promote greater acceptance of the foreign language program.

Discussion

Discussants: Carroll; Davies, Nickel, Sibayan, Tai, Worotamasikkhadit

In many South East Asian countries, English is an obligatory course, in some cases even at primary level, so that aptitude testing would be applicable only at higher levels of education; e.g. for students intending to specialize in English studies.

The various factors tested in the MLAT seem to be of equal predictive effectiveness, and a hierarchy among them would be hard to specify because they are all interrelated in a system. The fact that "grammatical sensitivity" seem to be an important factor in language aptitude would imply that students, possibly on a more advanced level, could profit from instruction about language structure. Experience shows that audio-lingual habit learning and cognitive-code learning are both desirable, in the language program.

The MLAT items testing grammatical sensitivity were not chosen on the basis of any particular theory of grammar, but for predictive value. The test is pre-TG, however, and if any grammar can be said to have guided it, it was that of Charles C. Fries. It might be noted, however, that Robert Lees examined the items at one time, and found nothing to object to.

Beyond aptitude there is the question of motivation. The motivated student puts more energy into his study and is more likely to persevere in it, hence achieve good results.

DISCRETE VERSUS NON-DISCRETE TESTING IN ENGLISH LANGUAGE PROFICIENCY

by EUGENE J. BRIERE

Over the past several years, different methods of teaching foreign languages or second languages have been used.

In the beginning, language was considered synonymous with literature. Put differently, language was thought to consist of the printed words contained in books — preferably in books which had been written by prestigious authors. Consequently, classroom teaching consisted primarily of learning the writing system of the target language and then reading literary passages in the target language or translating these passages into the native language. When any attention was paid to anything other than translated prose, school programs were designed to provide memorization of verb paradigms or parsing of written sentences.

The tests or evaluation procedures developed from these literary-grammar-translation methods of teaching consisted of compositions and dictations in the target language or grammar translation exercises. Clearly, the scoring, evaluation and grading used in the three techniques were subjective, and it was difficult, or sometimes impossible, to assess the students' resulting performances in any systematic objective manner. Stylistics, spelling and the examiners' personal prejudices frequently interfered with objective evaluation of achieved results and with reasonable predictions of success or failure in future learning in the target language. Since the variables of interest were not defined precisely, each examiner could use a different set of criteria for grading a composition or a translated passage. Some teachers placed more emphasis on grammatical precision (based of course on some literary style) while others were more concerned with imaginative, complex performances or stylistic considerations which showed "creativity" rather than simple-minded, grammatically precise inanities such as the "Look-Jane-Look" "See-the-ball" variety. Frequently, the net result was that fifteen, or more, different teachers could evaluate the same composition in fifteen, or more, different ways.

If any attention at all was given to developing oral proficiency in the target language, the ensuing "tests" frequently consisted of unstructured interviews or "oral compositions" which lead to the same chaotic conditions in evaluation as those described for the written tests.

After the structural linguists such as Fries and Lado began to emphasize the primacy of proficiency in oral language, teaching methods and testing procedures changed considerably.

For one thing, oral, structural pattern practice replaced the previous literary methods. "Discrete point" teaching and testing become the order of the day. What I mean by "discrete point" is the assumption that there are a number of specific things, the knowledge of which constitutes "knowing" a particular language and that these things could be precisely identified at the different levels of syntax, morphology and phonology. Lado[1] for example, using a paradigm developed by psychologists to identify pro-active interferences,[2] assumed that a contrastive analysis of the native language with the target language could precisely identify those learning problems which would be encountered by native speakers of L_1 attempting to learn a specific target language. Moreover, the learning problems identified through the procedure of contrastive analysis could be developed into discrete point teaching materials or tests by simply writing patterns or test items for each of the learning problems involved. In actual practice, however, the things which are thought to constitute a language are not frequently defined through contrastive analysis but more often through a structural analysis of the target language only. The ensuing identification of the phonemic contrasts, the morphemic privilege of occurrence in certain pattern slots, the vocabulary items and the contrasting sentence patterns which are to be taught and then tested are frequently chosen in a very arbitrary manner. TOEFL[3] is an example of a discrete item test which was definitely not based on contrastive analysis.

Perhaps the largest gain made in turning to discrete point testing was a specific identification of the categories to be tested and an objectivity in scoring which was impossible with the translation or composition type of tests. Multiple choice items can be statistically analyzed for difficulty scores, discrimination scores and r correlations with external or internal criteria. We now have a method of reducing reliability and validity to a number which we can readily and easily understand. However, there is a growing concern among certain language test designers over the *actual* validity of this discrete point approach because of the very difficult problem of identifying precisely many of the complex variables which define the competence of a speaker or listener in any act of communication.

There is a growing agreement among psycholinguists and sociolinguists that traditional linguistic definitions of the notion of competence in a language are too narrow and are inadequate in identifying all of the skills involved when two people communicate. Consequently, discrete item language tests based on the narrow definition of linguistic competence will be inadequate. At best, such tests only give us some kind of measure of behaviour which I will call "surface" behaviour based on the analogy of a floating iceberg. The part of the iceberg which is seen floating on top of the water is but a small fraction of what lies underneath the water. So it is with language competence. We suggest that the language tests being used today are limited to measuring that which is on the "surface" and can give us no information about what is "underneath". However, probably it is precisely these unidentified and unmeasured variables "underneath" which constitute the "bulk" of language competence. What is needed is a serious attempt to develop a model which will identify and measure those variables which, at the moment, are "underneath the surface". I'd like to spend the remainder of this paper briefly summarizing some of the serious and sophisticated attempts which are currently being conducted in the United States.

Bernard Spolsky,[4] University of New Mexico, suggested, in a recent paper, that although Fries rejected the layman's notion that knowing a certain number of words in a language constituted the criterion for knowing that language, he still maintained the related notion that knowing a language involves knowing a set of items. Spolsky suggests that testing of individual elements such as sound segments, sentence patterns, or lexical items, is still inadequate.

He points out that the layman's criterion for *knowing* a language is usually expressed in some type of *functional* statement. For example, "He knows enough Thai to read a newspaper and ask simple questions for directions". Statements such as these refer to language *use* and not to grammar or phonology. The question then arises, how does one go about deciding when someone knows enough language to carry out a specified function? One approach would be to give someone a language-using test to perform such as having a physics major listen to a lecture on thermo-dynamics and then test the comprehension. Another approach would be to characterize the linguistic knowledge which correlates with the functional ability. However, one of the fundamental reasons that this approach has not proved successful is that it fails to take into account the fact that language is redundant and that it is creative.

Redundance (part of the statistical theory of communication) is present in all natural languages since more units are used to convey a message than are theoretically needed. Spolsky has experimented with redundancy as a testing technique.

In his experiments, noise was added to messages on tapes and the tapes were played to native and non-native speakers. The non-native's inability to function with reduced redundancy suggested that the key thing missing was the richness of knowledge of probability on all levels, phonological, grammatical, lexical and semantic. At least two implications followed from these experiments. The first is that knowing a language involves the ability to understand a message with reduced redundancy. A model of understanding speech must then include the ability to make valid guesses about a certain percentage of omitted elements. The second implication is to raise some serious theoretical questions about the value of deciding a person knows a language because he knows certain items in the language. The principle of redundancy suggests that it will not be possible to demonstrate that any given language item is essential to successful communication, nor to establish the functional load of any given item in communication. He makes the distinction between language-like behaviour, for example, the utterances of a parrot, and knowing a language on the basis of creativeness, that is the ability to produce and understand a sentence which may never have been heard before. One fundamental factor involved in the speaker-hearer's performance is his knowledge of the grammar that determines an intrinsic connection of sound and meaning for each sentence. We refer to this knowledge (for the most part, obviously, unconscious knowledge) as the speaker-hearer's "Competence." Therefore, in searching for a test of overall proficiency, we must try to find some way to get beyond the limitation of testing a sample of surface features, and seek rather to tap underlying linguistic competence. Testing selected items can only give us a measure of surface behaviour or performance.

John Upshur,[5] University of Michigan, feels that attempts to measure a "general proficiency factor" have been essentially unsuccessful primarily because of the lack of any performance theory generally available to, and useful for, those who might prepare production tests. He has suggested a small scale model of an internal machine which would specify some of the variables needed in a performance theory.

Upshur suggests that in the act of communication a Speaker's Meaning (SM) be distinguished from Utterance Meaning (UM) or Word Meaning (WM). In communication it is the task of the producer (a task certainly shared by the receiver or audience) to "induce" in, or transmit to the audience a meaning (AM) which has as a part an equivalent of SM.

Because communication requires that AM contain SM, because UM is a medium through which this is accomplished, and because SM need not be equivalent, more is required than that S (and A) have competence in some language. For S to get his meaning across to A (i. e. to communicate, to have AM contain SM), it is necessary (1) for A to get the word meaning, (2) for A to know the case relations for each W (this seems to be a part of UM), and (3) for A to get the relations between a proposition and other concepts.

Upshur then develops a model for A which could account for the processing of SM in A. The kinds of "components" in A which he suggests are such things as: perceptions of the outer world (PO); a store of concepts (AMS) resulting from the current communication transaction (CCS); a semantic net (NET); a linguistic competence (COMP) and several others.

From S's point of view, he must have a concept to communicate (SM), and some reason for doing so. S has the belief that A lacks the concept SM, and cares to have it. His communication ability is then a function of (1) his success in determining the constraints imposed by the contents of A's components, (2) his success in altering the contents of those components, and (3) his success (in language communication) in adapting his own competence.

The model suggests that oral production testing, viewed as one of the four skill components of the 1961 Carroll model, is but one part of speaker communication testing. Communication measurement involves a matching of SM and AM, therefore "precise" measurement is not likely without comparable measures of both.

One experimental form to test one kind of communication situation has been and is being investigated by Upshur. In this technique, a set of 36 four-picture items was prepared. The S's specific task was to communicate to a remote A which one of the four pictures was identical to a single picture shown to him by an examiner. Students of English as a foreign language took the intial 36 item test and the utterances were recorded. Four native A's listened to the tapes. The inner judge reliability for correct items was .87. Uniformly high coefficients were found between raw scores, total response and communication rate scores with composition and achievement tests scores. (Incidentally, we[6] are currently using a modification of this technique to elicit oral responses on our project to develop ELS proficiency tests for North American Indian elementary school children).

Leon Jakobovits[7] from the Centre for Comparative Psycholinguistics at the University of Illinois points out that there is an obvious difference between linguistic competence as it is traditionally defined and communicative competence. The latter involves wider considerations of the communication act itself; considerations which the linguists have dismissed in their definitions of linguistic competence as being primarily the concern of paralinguistics, exolinguistics, sociolinguistics and psycholinguistics. Since the authors of language tests are aware that the study of language *use* must necessarily encompass the wider competencies involved in communication competence, the development of language tests must move from the present position of measuring merely linguistic competence to the position of measuring communicative competence.

Jakobovits points out that speakers of a language have a command of various codes that can be defined as a set of restriction rules that determine the choice of phonological, syntactic and lexical items in sentences. For example the choice of address form in English, "using the title Mr. followed by the last name versus first name," is determined by the social variable which relates the status relation between the speaker and the listener. These selection rules and others of this type are as necessary a part of the linguistic competence of the speaker as those with which are more familiar in syntax, such as accord in gender, number and tense; and it would seem to be entirely arbitrary to exclude them from a description that deals with linguistic competence.

In order to be able to account for the minimum range of linguistic phenomena in communicative competence, it will be necessary to incorporate in the analysis three levels of meaning, namely linguistic, implicit and implicative.

By "linguistic meaning" Jakobovits refers to the traditional concerns of linguists such as Chomsky and Katz. This includes a dictionary of lexical meanings and their projection rules, syntactic relations and phonological actualization rules.

By "implicit meaning" he refers to the elliptically derived conceptual event which an utterance represents. By this is meant that particular implications for homonymous utterances are a function of the situational contexts in which the utterance is used.

In order to recover the particular meanings of the word intended by the speaker, the listener must engage in an inferential process which makes use of his knowledge of the dictionary meaning of words as *well* as his knowledge of the overall situation to which the sentence as a whole refers.

"Implicative meaning" refers to the information in an utterance about the speaker himself, e.g. his intention, his psychological state, his definition of the interaction, etc. In some cases these implications are necessary to recover the intended meaning of the utterance. For example, "Do you have a match?" is not a question to be answered verbally, but a request for fire to light a cigarette.

The problem, then, of assessing language skills becomes the problem of describing the specific manner in which an individual functions at the three levels of meaning just identified. Language tests, then, must take into account the full range of phenomena if communicative competence of language *use* is to be tested.

Jakobovits makes some tentative suggestions with respect to some methodological approaches which may be used in connection with his classification scheme. Some of the methods suggested are as follows:

> *A. Judgements of Acceptability* — ask a subject to judge the acceptability of an utterance or pick the most appropriate of two similar utterances.
>
> *B. Semantic differential techniques* — subjects rate a word on a seven point bi-polar adjectival scale according to the Osgood method.
>
> *C. Acting out situations* — ask a subject how he would say something under specified conditions in order to assess his encoding skills in terms of the different kinds of meaning just described.

Even scholars in the field of neuro-physiological speech are beginning to question discrete point teaching, and presumably, discrete point testing. In a paper entitled "Physiological Responses to Different Modes of Feedback in Pronunciation Training", Richard Lee[8], Florida State University, reported some exciting experimental results to the TESOL conference in San Francisco two months ago.

Working on the premise that pattern drill and phonemic discrimination drill is *not* genuine language behaviour, has *no counterpart* in natural language behavior and produces boredom, lack of motivation and little learning among the students, he performed the following experiment.

It has been established that some physiological arousal is necessary for learning to occur. Arousal is most often measured by heart rate, galvanic skin response and breath rate. Of these measures, heart rate is the most robust.

Ten women and eight men, all foreign students taking ESL course, were measured on an E and M Physiograph Six. This machine is similar to the polygraph used in lie-detector tests but free from connecting wires to a central recording machine thus providing the student with complete mobility.

The heart rates of the students were measured at seven different points in time under two basic conditions: one was during normal conversation and the second was during pattern practice.

The peaks of arousal which are expected during normal conversations did occur. Unfortunately, the measurements during pattern practice showed such little arousal in heart rate that Professor Lee was led to believe that no learning was occurring at all. In fact, the lines during pattern practice were almost flat with a slight drop at the seventh or last reading in time.

Admittedly, this is a small population from which to extrapolate to the universe, but I certainly hope he continues with this rather unusual technique for measuring learning in hopes that we can gain some insight into the language teaching methods we are currently using.

The sociolinguistic works of Robert Cooper and Joshua Fishman at Yeshiva University, Charles Ferguson at Stanford and William Labov at Columbia (to name but a few) are providing language teachers, language testers and linguists with data which could lead to that "break-through" which is now needed if we are to move ahead into an era of sophisticated understanding of what to teach and what to test in order to provide psychologically sound understanding of the complex variables involved in communicative competence.

Furthermore, I would like to conclude with my personal bias which is that any *real* "break-throughs" and new insights *must* be the results of an interdisciplinary team of teachers, testers, psychologists, sociologists, linguists and many others. I can't think of any one single discipline (let alone a single person) which can provide all of the answers we now realize we need in order to provide new and exciting teaching materials and methods and truly valid test instruments to evaluate communicative competence.

References

1. Lado, Robert. *Linguistics Across Cultures.* U. of Michigan Press. 1966.
2. Jenkins, James, in Charles Osgood and Thomas Sebeok (eds.) *Psycholinguistics.* Indiana University Press, 1967.
3. "Test of English for Speakers of Foreign Languages." Educational Testing Services, Princeton, 1963 — on.
4. Spolsky, Bernard. "What does it mean to know a language, or how do you get someone to perform his competence." A paper presented to the USC Conference on Language Testing, 1968.
5. Upshur, John. "Measurement of Oral Communication." A paper presented to the USC Conference on Language Testing, 1968.
6. Briere, Eugene. "Testing ESL Skills Among American Indian Children" *Georgetown Monograph Series on Languages and Linguistics*, 22, 1969.
7. Jakobovits, Leon. "A Functional Approach to the Assessment of Language Skills" Paper presented to USC Conference on Testing, 1968.
8. Lee, Richard. "Physiological Responses to Different Modes of Feedback in Pronunciation Training." Paper delivered to TESOL conference, San Francisco, March, 1969.

Discussion

Discussants: Briere; Boonlua, Davies, Nickel, Ohtomo

The idea of definition of language learning objectives in functional terms — ability to do a task — is acceptable in general but difficult to make precise The combination of competence and skills, and the level of each, required to "read a newspaper," for instance, might be difficult to specify with sufficient precision to make the objective testable or even teachable, "Suitability of response" is difficult to use as a criterion because of the unpredictability of the suitable response in most situations, i.e. various responses may be "suitable." This would oblige us to consider, to some extent, probabilities of frequency, both in establishing test criteria and in developing some kind of hierarchy in materials for teaching. This, however, is not an easy task.

Although we aim at teaching language competence, we need a performance model of language learning. Experimental work in language learning must certainly deal with a performance model, and the language student, whose objective is language *use*, is really interested in a performance model. Yet we do not as yet have any adequate performance model, and when developed it may be quite different from the competence model.

In response to a query on why picture stimuli were used in the Navaho testing project, Dr. Briere stated that: (1) Written stimuli could not be used, because the students tested were generally at a low achievement level where reading was still a problem; (2) Experience had shown that the Navaho children often did not respond to direct oral stimuli — questions, greeting, etc., but *did* respond to the pictures as a kind of game; hence, the picture stimuli generated more response; (3) The pictures gave a factor of control which a simple interview did not.

TRANSLATION IN LANGUAGE TESTING

by HUYNH DINH TE

It seems an anachronism to talk about translation in this age of the language laboratory and teaching machines. In the last few decades significant innovations have been introduced into the field of language teaching and testing. I remember when a visiting professor came from overseas he was surprised to find translation used in teaching and testing, and said it was like using a bicycle in space exploration. So I am a little nervous in introducing a subject like this; perhaps even the title of this paper has aroused disapproval. But this Seminar is meant to review all the different techniques used in Southeast Asia. Translation has become obsolete in the eyes of many language teachers. And yet it is not irrelevant to bring up the problem of translation as a technique of language testing before this Seminar for, in spite of all those innovations in teaching techniques and the success of the language laboratory, translation is still largely used as a testing method in many countries of Europe and Asia, including the speaker's native country. The purpose of this paper is to reconsider the nature of translation and the extent to which it can be used as a technique of testing in the new context of language teaching of our age.

Since the end of the Second World War the general trend in language teaching and testing has been the emphasis on the monolingual approach. No matter what method is employed — the direct method, the active method, or the structural linguistic method — the emphasis is on the exclusive use of the target language and on the priority of speech over writing. The native language of the student is discarded. The meaning of words and sentences is explained by other words, visual aids or acting. It is not surprising that translation is rejected for it is mainly based on the written language and involves the use of the native language of the student.

The main argument brought out against the use of translation is that it is a difficult exercise requiring a high level of mastery of two languages — the native language and the target language — as well as a high degree of cultural and literary sensibility. In this sense translation is impractical as a method of language testing, especially at the beginning and intermediate stages of language learning. This argument is perfectly valid if we think of translation as an art and if the text for translation is a piece of literary creation of great value. One must have a perfect knowledge of the native language and the target language as well as a deep understanding of culture and literature in order to produce a good translation which does not betray the original. But there is a great difference between translation as an art for professional translators and translation as a device for testing the knowledge of students on certain points of grammatical and lexical structures. The text of translation used for testing should be carefully selected, structured, and graded so as to suit the level of the student. In many cases the text is composed by the teacher himself.

It is obvious that one must understand in order to translate but it is also true that one must translate to understand, chiefly at the beginning stage of foreign language learning. It may be said without fear of error that the learner cannot think of himself as having understood a word or a structure until he has found an equivalent word or structure in his native language. Consciously or unconsciously he will abstract translation equivalents for himself from the observation of the target language in operation.

Translation is also rejected because it is thought to hamper the process of language learning. It has been contended that the student should discard the native language to avoid its interference with the target language. Since it involves the use of the native language translation fosters this undesirable interference. Evidence of this interference can be found in the mistakes made by the students in their translation exercises. Instead of writing English the students produce English renderings of the grammatical structures and lexical peculiarities of their native language. Translation is also thought to hamper the mastery of one's mother tongue. It has often been pointed out that the native language has never been more awkwardly written than in exercises in translation from a foreign language.

The conflict between the grammatical and lexical structures of the languages involved is responsible for most errors in translation. On the grammatical level the target language and the native language may use different devices to express the same structural meaning or relationship. In this case the student is apt to overlook in the target language what is not grammatically significant in his native language structure. The two languages may also use the same grammatical device but in different ways. The native language and the target language may both use word order, for instance, as a syntactical device but the order of words may be different in these languages. Mistakes of the type, "horse race: race horse", are not uncommon for Vietnamese students making translations from their native language into English.

On the lexical level the main source of error in translation arises from the difference in semantic areas covered by a word in the native language and in the target language. Practically, there is almost no one-to-one correspondence in the meaning of words in different languages. Moreover, figurative meanings and connotations of words seldom remain the same from one language to another. Sometimes words in one language have no equivalent in another language. Mistranslations also come from the difference of semantic units in different languages. A single word in one language may correspond to a group of words in another. Translation equivalence is contextual rather than formal.

It is a fallacy to think that we can avoid the interference of the native language by avoiding its use in second language learning. With or without translation this interference will work on the unconscious level by the brain mechanism itself. Unlike the child who learns his native tongue, the adult learner of a second language has a built-in system of linguistic resistance. The problem is certainly not to avoid it but to overcome it. Translation is a way to help the student realize the structural conflicts between the foreign language and the native language.

Translation as a method of testing has certain limitations. For obvious reasons it cannot be used when the teacher does not speak the native language of the students or when the students do not have the same native language background.

This limitation explains why translation is neglected in the teaching of English as a second language in English speaking countries such as England or the United States. Translation can only be used when both teachers and students speak the same native tongue, in countries where English is not the native language.

Obviously translation cannot be used to test the proficiency of a student in English pronunciation, his ability to express himself in oral English, or his skill in creative writing. This is why translation is not a popular technique of language testing nowadays when emphasis is laid on the oral-aural approach and the main objective in English learning is to acquire the skill of understanding and speaking the oral language.

Another disadvantage of translation as a technique of testing is that it cannot always be scored with objectivity. Much depends on the personal taste and preference of the teacher. In many cases the same sentence can be translated in different ways into another language. Sometimes teachers do not agree on which is the best translation of a given text. In terms of the language learning process translation is primarily based on the stage of selection, which is the most advanced stage in language-habit formation. The student has to select the single appropriate equivalent among the various possibilities offered him in the target language or the native language.

Translation is used to measure the student's ability to manipulate the various grammatical and lexical structures of the target language. The two different kinds of translation — translation from and translation into the target language — will serve two different purposes. The former is used to test the student's ability in recognizing form, structure, and meaning in the target language; the latter is used to test his ability in manipulating the target language on the grammatical and lexical levels of structure.

Much of the success of translation as a technique of testing depends on the choice of the text for translation. The text should suit the level of the student and contain only structures and lexical items already learned. Long and unselected passages containing difficult and unfamiliar words or structures only cause discouragement to the student and disappointment to the teacher. Unfortunately, translation usually means this kind of translation in which the student is confronted with words and structures he has never seen before. This kind of practice can only give discredit to translation as a method of language testing. Difficult and unselected texts should be reserved for professional translators or, at least, very advanced students who have more or less mastered the target language.

At an advanced stage of language learning, translation is an appropriate technique to test the student's ability to select the most accurate term among several words expressing various shades of meaning in the target language. It is also useful to measure the student's ability in selecting the right meaning among the various semantic possibilities of a term in the target language.

For intermediate and beginning students the text for translation should consist of carefully selected sentences which contain only structures and lexical items previously learned. The purpose of translation as a testing technique is to elicit from the student linguistic responses which will show whether or not he has mastered a grammatical pattern or a lexical item already learned. It does not aim at gauging the creative power or the puzzle-solving ability of the student. In this sense translation is neither better nor worse than other methods of testing based on the monolingual approach such as the true-false, multiple-choice, substitution, and completion tests.

It is used as a stimulus to elicit from the student a controlled response which will allow the teacher to measure the ability of the student in the target language. The main thing for the teacher is to select the various features of the language structure which will be tested and devise materials which will suit the level of the student.

Since the purpose of testing is to measure the achievement of the student in the target language any technique which fulfils this goal should be used freely without fear. As each technique can test only a few facets of the student's ability and a few levels of the target language structure, a comprehensive test should include several techniques based both on the monolingual and bilingual approaches, such as objective tests, essay writing, and translation.

Discussion

Discussants: Te; Lee, Mohan, Ng, Nickel, Otter, Phillips, Sculthorpe

It is significant that Te's brief for the use of translation in language teaching and testing received quite general support. It was pointed out that, even though a teaching methodology may rigorously eschew translation, the pupil's learning strategies involve considerable translation, either consciously or unconsciously. Only at quite high levels of language learning is complete bilinguality approached; doubt was expressed whether complete bilinguality was *ever* achieved. The L1 is a barrier through which the student must pass to reach the L2, but it is also a means by which he reaches the L2.

Translation as a testing device must be distinguished from translation as a teaching device, but this distinction need not be too sharp. What is testable should have a foundation in teaching. If translation is to be used as a testing device, there is needed some specification of standards; basically what is desired is functional equivalence, not exact equivalence. As a testing device, translation offers the advantage of being more accurately scorable than free writing, and for that reason has been favored by test developers seeking objectivity.

Translation as a teaching device should not be ignored. It is a necessary component of the teaching/learning process. Bad teaching makes translation the *entire* learning technique, but this is no reason for eliminating it altogether. In language learning we need to learn *about* language, and translation is a good technique for developing awareness of the workings of languages and insight into language system. Such knowledge is particularly useful and effective on the more advanced levels of instruction, especially university levels.

Also, there is a need for capable translator–interpreters in many parts of the world, and one objective of language teaching should be to develop this capability. It is necessary to determine at what level one can feasibly begin instruction in the specific studies of translation, and to develop effective techniques for teaching translation. Development of a cadre of qualified translators/interpreters is worth attention, and is not an easy task. It is especially important to develop cross cultural awareness as part of the translation capability; in fact, more attention to this element should be given even on early levels of language learning.

ENGLISH LANGUAGE TESTING TARGETS AND TECHNIQUES[1]

by S. N. SAHA

1. Introduction

Considering the profusion of literature in the field of language teaching with its emphasis on new instructional objectives, new instructional materials and new teaching methods, information available in language testing and evaluation is relatively meagre. While latest methods of teaching have brought about great improvement in most areas of language instruction, English language testing has not comparably progressed to fulfil its function as an important part of the teaching-learning process.

It needs hardly any emphasis that the central problem of all educational endeavour is learning. As in language teaching, the greatest immediate pay-off is determined to the extent that learning takes place in students, so also in language testing, the basic criterion should be that the tests are the truest reflection of the teacher's pedagogical objectives and of the students' expected level of achievement. So the language test should not be a collection of fragmentary facts about structure, vocabulary, etc., but should be a purposeful, stimulating and rewarding experience for the students as they exercise their skill of putting together the various elements of language into a fabric of habits. It is therefore of the utmost importance that the test serves to help, and not to hinder, the progress that the students make in language. Improved evaluation is thus the core of a programme to bring about the qualitative improvement of teaching and learning.

2. Limitations of English Language Testing

Unfortunately, most current examinations are narrow and uninspired and do not exercise a conducive influence on language learning. Even a cursory survey of the English Examination Question Papers will pointedly show that they consist mainly of a certain amount of predictable grammar, and that again without reference to context, and straight recall of factual information gleaned from the texts. Similar types of question are repeated each year with the result that students concentrate on intensive preparation in relatively few areas and the teacher complacently encourages them to persist in this effort. The topics for composition are chosen from such a limited, stereotyped range that the students' answers do not bear evidence of the extent of their use and control over the language, unspoilt by the crammer's art. Essay questions on the prescribed texts do not at all stimulate the student to select and integrate or to actively reconstitute the learnt material. Short answer questions on the prescribed texts, supposed to test comprehension, are nothing more than mere recall, thus putting a heavy premium on the students' memory, and consequently on cramming. Again, questions on an unseen passage are set so sequentially that whatever effort at comprehension the student has to

---

1. This paper reflects the thinking that has been developed in the area of English language testing in the National Institute of Education of the National Council of Educational Research and Training under the Ministry of Education and Youth Services, Government of India, as a result of the intensive work for the improvement of English language teaching and testing in the different states in India.

exercise reduces itself to merely locating the sequence in the passage. Understandably such an ill-assorted, stereotyped pattern of questions does not obviously constitute the corpus of testable elements nor does it to any reasonable extent reflect the students' ability to read and write English. Moreover, such a pattern of questions is not linguistically and pedagogically sound. In terms of language it is not based on carefully sampling a wider range of material amenable to testing at one sitting, while pedagogically it neither motivates and directs student learning in desirable ways nor encourages interesting and stimulating work in the best traditions of language teaching. So the vital question of language and the learner is lost sight of, so much so that the language paper inevitably becomes more error-oriented than success-oriented, just seeking evidence of pupils' deficiencies, and not of their ability to effectively use the language in real-life situations. Fortunately during recent years, many examining bodies are becoming increasingly alive to the limitations and shortcomings of English language testing and, as a result, a welcome shift of emphasis is under way.

3. **Language areas to be tested**

Desirably a test should secure a more effective measurement of language skills; i.e., it should help to give a true picture about a student's familiarity with the different aspects of the language, his basic knowledge, understanding and skill. It should, therefore, be comprehensive so that it could serve as a means of determining the status of the learner in each learning area. This presupposes that the teacher or examiner should have a clear conception of the language areas that may be tested and is interested in determining whether the student has reached a satisfactory degree of achievement in these areas.

While considering the areas and the problems to be tested, it is in order to consider that language skills remain highly complex because their elements are systematically interlocking and interactive and hardly ever function independently of each other. So a choice has to be made between language as an integrated stream of spoken or written expression with all the elements contributing to it and as separate elements like pronunciation, grammatical structure, vocabulary, spelling, etc., since all these areas are highly interrelated and over-lapping. Indeed, both have their well-deserved place in the scheme of language testing, since the testing of one alone will not give a very authentic picture. So, whether language is tested as an integrated skill with all the elements woven into it, or whether it is tested in terms of its separate elements, there is no denying that it is language that is being tested. What counts as more important is the advance decision on the objective that the teacher or the examiner has to make, while preparing test items. If it is merely to diagnose spelling difficulties of students, then a test with a fairly large range assessing the students' performance on, say, writing words from dictation or judging whether a particular word is correctly or incorrectly spelled, is undoubtedly useful. But if the objective is to test whether the student can show the same spelling ability in writing a composition, isolated spelling test scores may not be adequate. In the same way a student's knowledge of grammar, his ability to recognize a sentence pattern, or even his ability to perceive a logical relationship between two or more things may be tested as isolated skills, but that does not guarantee that his ability to integrate these various skills is amply evident in speech or writing. To be optimally effective and fairly comprehensive, a test should desirably aim at evaluating these skills not only in isolation but also in actual practice. The teacher and the test constructor therefore must be very clear about the purpose of each question he constructs.

What specific teaching-learning objective is involved? What specific learning outcomes are expected? What evidence are students expected to give of their assimilation of language material? What will be the proportion of questions testing the students' skill in the use of English, of questions measuring their skill in reading comprehension, of questions testing their knowledge of the fundamentals of language and of questions testing knowledge of the subject matter? All these have to be considered in the light of broad-based sampling of the functional elements of language, contributing to total skills. So it is desirable to devise some ways of sampling the full range of elements that make up the language and see that these elements are represented in terms of their relative importance in order to secure effective communication. In this connexion it is well worth mentioning that different types of evaluative procedure could be pressed into use, such as, multiple choice, short answer and essay, to sample a wider range of language material in as wide and varied a spectrum of contexts and situations as the time allows.

4. Testing Various Skills

Evaluation of learning is a complex and cumulative process of determining whether objectives of learning are being or have been achieved. It is more so in respect of the examining of English language which is more widely taught for more practical and utilitarian reasons than any other language. Proficiency in one skill, say, reading, does not necessarily mean equal proficiency in skills such as oral expression and listening comprehension. Various skills, therefore, should be kept in perspective, and, accordingly, the examination should be so constructed that these skills are suitably evaluated, and no skill is neglected at the cost of others.

5. Testing Listening Comprehension

Listening comprehension is one of the most important acquired skills. Yet in marked contrast to reading, it is the one in which the student usually is the weakest. In the light of the great importance of listening as a learning device, its inclusion as one of the measureable skills cannot be too strongly emphasized.

Listening comprehension may form part of either oral or written tests. Identification of sounds and differentiation in sounds (as in late/let), stress differentiation in word-groups (one armed bandit/one-armed bandit), stress differentiation in sentences (*What* are you doing/What *are* you doing?), intonations to indicate question, statement, exclamation, hesitation, uncertainty, use of contextual clues to word meanings, recall of important details presented orally, identification of the central idea of the statement given orally, recognition of sequence signals or transitional elements in sentences, identification of the various uses of language in different spheres of activity, etc., are some of the test elements aimed at assessing the student's skill and acuity in listening. The potential of listening comprehension as a learning instrument will be defeated if no testing instruments and procedures are devised to measure this skill. However, even without them, and without its being a testable area in the public examination, resourceful teachers may reap rich dividends by improvising effective contexts and situations in class.

6. Testing Oral Expression

Ability to speak is one of the most highly prized language skills. It is claimed that oral proficiency makes reading and writing in a second language easier to learn. Yet it is the least developed and the least measured of all language skills. Any examination which does not include a test of the candidate's ability to express himself orally, is incomplete. A written examination with all its attributes cannot, by itself, claim to be an adequate measure of the candidate's ability to express himself orally in life situations, or the ability to report acts or situations in precise words, or the ability to converse, or the ability to express a sequence of ideas with ease and fluency. However, there are very considerable practical problems in trying to measure skills in oral expression, very different from those normally associated with reading and writing.

The following practical problems merit careful consideration:

Determination of outcomes to be tested in oral expression.
Determination of test material to be used for oral expression.
Formulation and specifications of criteria under different dimensions.
Procedures of conducting tests of oral expression.

6.1 **Determination of outcomes**

It is desirable that those outcomes which come under the purview of written tests should not be included in the tests for oral expression. However, it is rather difficult to isolate outcomes which are absolutely specific. For example, organisation is as much a quality of oral expression as of written expression. The case is similar with vocabulary and grammatical structures. It is, therefore, very necessary to determine in advance, while planning tests for oral expression, those outcomes which may receive greater emphasis and those in which emphasis may be subdued.

6.2 **Test Material**

Test material should be based on the immediate environment of the students. It may include:

a brief conversation by way of rapport,
specific short answer questions and
a sustained speech, calling upon the student to talk without interruption for two or three minutes on a very familiar topic.

The first one may be aimed at assessing manners and etiquette, reciprocity and flexibility. The second may be designed to assess in essentially normal communication situations such aspects as pronunciation, stress, intonation, grammatical structures and vocabulary at a normal rate of delivery, and the third one to assess fluency in expression.

The test of oral expression should not be a prolonged one; it should preferably be limited to ten to fifteen minutes.

6.3. **Formulation and Specifications of Criteria under different dimensions**

The marking of the attributes or aspects of spoken English is fraught with the problems similar to those of the marking of the qualities of written English; indeed in the former case they are more pronounçed. While the evidence of the written composition may be evaluated at leisure, re-scanned and reassessed, that of spoken English is fleeting, must be caught at the moment of utterance and is not amenable to reassessment by another person. It is rightly said that even a recording used for repeated playback cannot include the full context in which significant utterances occur. Again, assessment may be highly subjective and impressionistic to the extent that one examiner will notice vocabulary, another fluency, still another, personality traits and so forth.

To make tests for oral expression a really effective and worthwhile measurement in the total scheme of examination, it is imperative to determine what weightage, maximal and minimal, may be allotted to different dimensions.

Each dimension has to be further clarified in terms of specific criteria to guard against inter-examiner and intra-examiner variability and thereby to ensure greater uniformity.

For example, what is meant by fluency in a three point rating — A, B and C — has to be stated precisely.

A. Speaks with facility and ease.

B. Occasionally halts as he speaks in order to search for the right word or to correct an error.

C. Speaks without confidence and so haltingly that it is difficult to follow what he says.

Again, a sustained speech may have the following criteria:

A. Speaks completely and correctly, though there may be some errors of vocabulary, grammar, pronunciation, which have little or no effect on the understandability of the speech.

B. Speaks with more errors, yet it could be understood.

C. Speaks in such a manner, grammatically, lexically and organisationally, that it conveys no sense at all.

6.4. **Procedures of Conducting tests**

Since the evaluation of oral expression is highly subjective, it is desirable that there should be three examiners. They should be given specific instructions as regards the procedure of approaching the response. If the item tests a grammatical problem depending on word order, the examiner may be instructed to disregard other qualities of oral expression, say, pronunciation and fluency. If again, vocabulary is to be tested, the examiner should disregard everything else.

Much has been done towards improving written examinations. Comparatively little has been attempted in the testing of oral expression. In view of some practical problems and of the difficulty of isolating discrete speech skills from the speech as intergrated skills, oral testing on a wide scale has not formed a routine part of the testing programme. However, that should not prevent the well-meaning educationists from paying increased attention to oral expression, continuing experiments in the construction and administration of oral tests and devising suitable techniques so that in the foreseeable future oral testing may occupy an increasingly important place in the English language examinations. Until then, oral testing should desirably receive increased emphasis in internal evaluation.

7. **Testing structure and vocabulary**

Language tests should assess as many facets as possible of language ability. It is therefore essential to test separately the elements of language as a representative sample of the learning problems, making sure that the student has adequately mastered the language. In testing knowledge of the grammar of the language, care should be taken to see that the test does not aim at a knowledge of the terminology, since the knowledge of the terminology is no indicator of how much of the language a student knows. A test should aim at assessing the control which the student is able to exercise over the grammatical structures of the language, and as such it should cover a wide range of the elements of language which basically contribute to any act of communication.

Tests of vocabulary, as are in vogue, have sought to elicit answers in the shape of definitions and explanations, which in all conscience, are not valid tests of language, since the situations themselves are not linguistically valid. It is only in context that vocabulary can be properly tested, for it is the context that appropriately determines the meaning of a word. In a vocabulary test what counts most is the power of discrimination, the sense of appropriateness, that gives the feeling for the language and helps to determine the meaning of a given word or words in contextualized situations.

In the area of vocabulary also, there is the need for a reasonable sample. Though it is very difficult to define with any degree of exactitude the range of vocabulary likely to be valid and reliable for this test, it is commonly held that the vocabulary test should be comprehensive enough so as to give adequate evidence of the student's ability to read with understanding articles of general interest in books and periodicals. Objective and very short answer questions should be set to test a wide range of structures and vocabulary which the students need to use in their familiar fields of knowledge and activity.

Tests of vocabulary and structure should not be based on fragmentary facts, unrelated to context. They should be properly contextualized, since it is only by assimilating and using the language in context, and by doing so often enough, that students can become truly proficient. Moreover, quite apart from their educational value as tests, such contextualized tests of vocabulary and structure have a most beneficial feedback effect on both teaching and learning. In fact, testing and teaching are not mutually exclusive; they are intrinsically related parts of the total educational process.

8. **Testing reading comprehension**

Comprehension is best shown by the pupil's responses to new (previously unseen) material because the presentation of the familiar material may elicit only memorized responses without demonstrating real understanding. The questions on comprehension are expected to stimulate the pupil to give evidence of behaviour higher than recall or recognition of the plain sense of the passages, such as locating the title or the central idea, giving implied meanings of the passage, inferring the author's or the character's mood, finding the lexical and figurative meaning of words in context, etc. Tests of comprehension should preferably have three previously unseen passages, each of a different kind of writing: dialogue or conversation, narrative and description: to ensure that the student has developed the ability to read and comprehend different kinds of English prose. Such passages should be self-contained, and the questions asked should strictly be limited to the content of those passages.

Ideally, comprehension should be separated from expression. Many students are unable to express themselves but they do comprehend the subject matter fairly well. In questions on comprehension the examiner is, strictly speaking, expected to give marks only for comprehension, but consciously or unconsciously, he is prejudiced by poor expression and holds back the marks the student legitimately deserves. It is, therefore, desirable to set only objective type and very short answer questions to do justice to the students' genuine skill in comprehension.

In this connexion it is relevant to point out that the traditional practice is to set a few questions on the prescribed texts. These questions ostensively aim at testing the student's skill in comprehension and expression, but in reality they test neither. Most such questions, in the case of students with a good memory elicit mugged-up answers, reproduced verbatim either from the texts or from the help-books; for others less fortunate, such questions spell disaster, and it is for the examiner to locate some relevant clues in their answers to justify his awarding at least some marks as a good gesture.

In case such tradition-bound textual questions are unavoidable, it is desirable to evolve a strategy so that such questions do not solely require memorization nor seek to elicit expression which is far from the student's own. To obviate undue dependence on memory, and therefore on cramming, relevant passages and stanzas from the texts may be quoted on the body of the question paper, and short comprehension questions may be set on them, to answer which students may not just require to depend on the resources of recall. Again, if expression is desired to be tested through the medium of known subject matter, it is better, and perhaps pedagogically sound, to give some relevant points and ask the students to organize them into one or two paragraphs. Such questions are virtually impossible to cram for.

9. **Testing Composition**

Free composition is not a satisfactory teaching device for students who offer English as a second language. It leaves them free to transfer to English the vocabulary and sentence structures of their mother tongue and, as such, it reinforces these incorrect habits. Asking students to try their hand at free composition will influence classroom teaching. Hence it loses sight of the objective of the composition class which is aimed at teaching a specific kind of writing with its own characteristics of vocabulary, sentence structure and paragraph arrangement. A student of English as a second language needs as much help in learning to write English as he does in learning to speak English so that he can, with ease and facility, reduce the structures he uses most frequently in writing to habit. So questions should be set on some functional, contextualized topics, each being a specific kind of writing, to test the students' skill in composition and the extent of their control over the language. Such controlled, contextualized tests, which do not just leave the students in a vacuum but help them to come out "with something to say" have potential feedback value in classroom teaching. So far as the test of letter-writing is concerned, it may take the form of a letter (only the communication of message) printed on the question paper, to provoke the appropriate response to it and to test skill in letter-writing. The questions on composition to be thus set should be definite, real and purposeful with regard to both teaching and testing.

It should be mentioned here that a single composition still tends to be a limited exercise as a test in writing. With ingenuity, a student is often able to avoid betraying the gaps in his knowledge by using only those words and structures that he thinks he is confident of. To ensure greater validity, it could be supplemented by more specific tests in controlled composition, within the same over-all time, to measure the various aspects of student progress towards the acquisition of writing skill. Since creative ability is not primarily looked for in tests of composition in a second language, it is desirable to provide the students with the necessary information.

10. Planning the Test

It is of vital importance to consider the planning of a test. A good test does not just happen, it is the result of careful planning. Moreover, planning must ensure that a test has balance and comprehensiveness. Such planning is imperative whether it is a question paper to be set for an external examination or whether a teacher is giving an examination to cover the work of a year, or a term in school, or a unit of a lesson after its completion, since there is the question of ascertaining the scope and type of mastery the student has achieved after completing a certain unit or course.

Given a particular topic or a unit or the syllabus to be tested at a specified grade level, the following ten dimensions that mark optimum contribution to the planning and make-up of a test may be considered:

Objectives to be tested.
Content to be covered.
Forms of questions to be used.
Situations to be organized.
Number of questions to be included within the specific time limit.
Level and range of difficulty of questions to be determined.
Scheme of options to be decided upon.
Language of the question to be made precise.
Items to be written, edited and assembled.
Scheme of scoring to be formulated.

10.1. Objectives to be tested

It is very necessary to clearly envision the objectives of instruction to be tested, and define them each in terms of specific behavioural manifestations expected of the students. Such specifications facilitate construction of items and also help the test constructor to see how far his test shapes itself in respect of balance and comprehensiveness and is intended to appraise the extent to which the student has developed in the desired ways. It is also necessary to determine whether the task required of the student in answering the questions pertains, in fact, to the intended objective. If the intended objective is 'comprehension', the test constructor is to make sure that the questions do not elicit any memorized responses. Trivial objectives may invariably be avoided.

10.2. Content to be covered

A test to be a reliable evidence of students' achievement, requires wide coverage of course content so as to obviate over-and under-emphasis on language areas in both teaching and learning. This necessitates a thorough analysis of content areas in terms of testable components. Although examination is always a sampling process, and each and every point discussed by the teacher cannot be tested, even ideally, still no effort should be spared to make the test reasonably representative. Decisions about objectives and content areas being made, it is necessary to prepare a two-dimensional chart in which both the objectives and the content areas are listed in such a way that the intended coverage of both may be clearly indicated. As one expert in test construction has said: "Such a chart of specifications should be prepared with great care before any item writing is undertaken. This chart is a blueprint for the test or examination and should be followed as closely as the workmen follow the architect's plan when constructing a fine building". As regards the weightage to objectives and content areas, it depends upon the relative importance of each objective in relation to other objectives and of each area of content in relation to other areas. Much discretion and judgement on the part of the test constructor will be brought to bear on it, since a carefully developed test plan serves to check the tendency to overload the test with items assessing knowledge of isolated facts or ephemeral details to the utter disregard of the more complex learning outcomes.

10.3. **Forms of questions to be used**

The decision about the form of the evaluation instrument will be determined by the objectives. If the purpose is to determine whether pupils can write well, i.e., can organize ideas, have fluency of expression, etc., there is no substitute for the essay form of the test. If the goal is to get evidence whether students are able to make discriminative judgement among items which are not too obviously alike or different, the multiple choice form of item may be an effective instrument. Short answer questions may be set to test more knowledge of facts or comprehension involving higher abilities. There is no one best form of item for all purposes, and hence there must be a judicious selection of appropriate forms for evaluation of particular objectives. If a test is aimed at evaluating progress towards several objectives and a wider range of language elements, it may comprise all or any combination of the types of items as discussed.

10.4. **Situations to be organized**

Language is behaviour, and behaviour can be learned only by inducing the students to behave. "When language is in action, there is always a speaker. He is always somewhere, speaking to someone, about something" (Nelson Brooks: *Language and Language Learning*). This is what has reference to the desirability of learning to make responses in situations which simulate real-life communication situations as closely as possible. It is, therefore, desirable that the test constructor should organize meaningful, connected situations in which students might feel that they are personally involved in meaningful communication within the limits of their control of structures and vocabulary.

10.5. **Number of questions to be included**

The total number of items to be included in a test will depend upon (i) the duration of the test and (ii) the nature and the form of test items to be used.

10.6. **Level and range of difficulty of questions to be determined**

Adjusting the level and range of difficulty of questions in the test requires proper judgement on the part of the test constructor. The difficulty level must be appropriate for the students to be tested and for the purpose of the item.

The range of difficulty may be such as will provide suitable opportunity for both the bright and the weak in the class. Difficulty level must be judged in advance, in terms of actual teaching experience with students at the educational stage being tested. Item analysis helps the test constructor gain a better understanding of how difficult an item is and how discriminating it is between those who have learned and those who have not.

10.7. **Scheme of option to be decided upon**

From the standpoint of a sound and dependable evaluation, the use of optional questions cannot be justified except for the desirable effect on pupil morale, since an examination with optional questions does not set the same task for all students, and each student demonstrates the achievement of different learning outcomes. Again, it tends to have an undesirable influence on study habits, since intensive preparation in relatively few areas is encouraged. If options are to be retained at all, they may be retained in essay questions, with the caution that optional questions are balanced and equivalent in respect of the objective, the content area, the difficulty level, and the form of question.

10.8. **Language of the question to be made precise**

The item should clearly set the task of the student in such a way that he will not make wholly irrelevant responses, even if he does not know the correct answer. A poorly worded, incomprehensible, or vague question defeats the purpose of the test. In such a case, the student most often omits the question, or tries to guess the intent of the question, or resorts to a random choice of answer. Directions for answering specific questions should also be clear and precise.

10.9. **Items to be written, edited and assembled**

Now comes the crucial question of writing individual items and editing and assembling them. The writing of items should bear upon this fundamental axiom: "The best test is one which is most valid, most appropriate to the levels, easiest to use, and has the most desirable effect on learning."

The writing of items is a highly specialized activity which is in part a personal gift and in part a learnable technique.

While setting essay questions, the test constructor should make sure that the questions have been precisely and unambiguously worded and do not have different meanings for different students.

The wrong choices, or distractors, in respect of a multiple choice item should be set with great care and imagination, and here advance planning is helpful. Each distractor, if it is to function in an item, must be one which some students will choose as correct. The best distractors are those that stem from the ignorance, misinformation and partial information that are likely to be found in students' responses, and experienced teachers are quite familiar with them.

Once the proposed items are drawn, they may be reviewed by the item writer himself, by at least one other teacher in the same field, and finally by an evaluation expert experienced in achievement testing. It is desired that test items reflect the confluence of teaching and testing.

A test again is more than a mere sequence of items, even if each one of them is fully valid in itself. Items of the same type should be grouped together and presented under a single set of directions. Usually they are arranged in order of difficulty, with the easiest first. Repetition must be avoided in order that students do not waste time and effort on items which test the same thing. The test, as it emerges from final editing and arrangement, should be such an organic whole that it affords ample opportunities to the students to display a totality of interwoven knowledge and skills.

10.10. **Scheme of Scoring to be Formulated**

Last but not least, in the total process of developing test material comes the scheme of scoring. In case of essay questions, the test constructor should furnish comprehensive outline answers and a detailed marking scheme to ensure uniformity of marking by different examiners. These will include all allowable alternative answers, the number of marks to be given for each part of the answer and the direction as to whether marks are to be given or deducted for different elements of the language, such as spelling, punctuation, grammar, etc. A scoring key is to be prepared for the objective type items. Here also, if there is more than one acceptable answer, that also has to be mentioned in the key. This entire process ensures objectivity in evaluation, consideration of weightage to different items, and detection of errors, if any, in the test. It is also worthwhile to have a final check on the test with an analysis to ensure that it has not marked a departure from the pre-determined purpose.

The foregoing outline may serve as a stimulus and direction to teachers and to test constructors who want to improve their competence in the preparation of test items. At all events, here is an exciting and challenging field of endeavour for resourceful teachers.

11. **Building up an Affective Base for Language Testing**

"Tests are only a means to an end. Tests as hurdles, as elimination devices, as bases for grades, interest me very little. But testing which teaches us how to teach, how to counsel wisely and effectively, how to understand the process of learning, how, in short, to help us assure that the human individual becomes what he can become — that is worthwhile, that is everything." What could be a nobler tribute to the educational value of testing than the above statement by an eminent educationist. What is said above of tests in general is very true of language testing, since learning the language and taking a test in it should both be stimulating experiences for children. Both have to be approached with the heart and with love. It has now become abundantly clear that how a student feels becomes more important than what he knows, and it is more so in a second language. So in teaching as also in testing, there should be more and more opportunities for students to gain feelings of use and control over the language. To the extent it is possible, the language test should not exercise a negative influence on students. Fear and anxiety should not be exploited as motivating devices. Every test that the student takes should make him feel better, more equipped, more able to cope with the challenge of language and more courageous and confident. So the language test should be so constructed that it attracts the learner instead of repelling him.

As a step towards fostering in students a positive attitude towards the language, there should be systematic review exercises with appropriate concentrated attention on various aspects of language before the final test is given at the end of the year or at the end of the school course. These review exercises should be developed sequentially and should be related more to success than to failure so as not to cause anxiety even among pupils near the lower end of the continuum of achievement. Such review exercises should bear on smaller units of learning, involving a week or two of learning activity. Each such review exercise based on a specific unit of learning is popularly known as a unit test or a diagnosis-process test. It is used mainly to determine whether or not a student has mastered the unit and what, if anything, the student has still to do to master it. In the case of the learning units which are basic and pre-requisite for many other units of the course, such as structures, tenses, vocabulary, etc., the tests should be frequent enough to ensure thorough mastery of such learning material.

Such unit tests, taken at frequent intervals, pace the learning of students, help motivate them to put forth the necessary effort at the appropriate time and foster their self-confidence, thus enabling them to take the final test without much emotional disturbance and unusual anxiety. They form an important part of the learning process and should have nothing to do with the judgement of the capabilities of the student or with the grading system. The basic advantage is that the student takes the test in a spirit of joy and adventure and begins to love the subject and to desire more of it.

12. **Conclusion**

Learning a language and articulating it is an emotional satisfaction. Anything to the contrary means frustration or perhaps a goodbye to the language. It is not an important question whether some students will adequately learn, quite a few will fail or get by and again some will learn but not well enough. What is more important is that most students can learn. In this International Educational Year it should be the effort of all English Language teachers and test constructors to think of the logistics of language testing so that it outgrows its limited purpose and sets off in a new direction to become a richer and a more rewarding experience for the students. Testing by itself is grossly negative in scope. Only when it functions as an ally of teaching is its significance seen in true perspective.

SOME PROBLEMS OF VALIDITY IN LANGUAGE TESTING

by P.W.J. NABABAN

1. **Introduction**

In teaching, as in other activities, it is important that from time to time an evaluation is made of the work for various purposes, e.g. in order to obtain information about the effectiveness of the teaching, the factors that need more emphasis, the need that may arise for regrouping the students, to predict performance in a higher course, etc. The evaluation is done through what we call tests or examinations.

Testing is an integral part of teaching; it is the means of feedback for the teacher by which he can improve his techniques and from which he may get pointers for raising the level of success of the curriculum. Testing is also used as a means of selection of incoming students and as a final evaluation of the students' success in following the program.

In order to ensure that the tests will be effective, they in turn need to be evaluated according to several criteria. The following are among the most important questions that should be asked about the test:

(*a*) Does it satisfy the purpose of the test? If a measurement of the skill in the manipulation of structures is the purpose of the test, then we have to make sure that it is that skill which is tested. We have to make sure that extraneous factors do not confuse the test items.

(*b*) Is the content of the test in accordance with the purpose and within the scope of the test?

(*c*) Do the techniques, or the situations, used to test control of the content of the test really do the job? Consider for example the example given by Lado[1]: If it is desired to test the auditory discrimination of the student between the phonemes / č / and / š /, the correct choice between the sentences *He is watching the baby* and *He is washing the car* will not necessarily indicate that the student hears the distinction between the two phonemes, because he would presumably give the correct answer, i.e. choose the correct picture, if he only knows the words *car* and *baby*.

(*d*) Is the test administrable? Can the test be given to the students considering their number and other circumstances. If equipment is required, is it available?

(*e*) Is the test scorable? Can the test be broken down into parts that can be scored conveniently? Are the scores of the test when given at another time to the same person relatively similar? Are the scores obtained meaningful? Do they reflect the degree of control of the things being tested?

(*f*) Does the test possess an extra usefulness? Does it have a backwash effect for the syllabus, curriculum and/or teaching methods?

Out of these questions, the first three are the most important from a technical point of view. They are concerned with relevance, and the common technically worded question would read: Is this test valid? We will return to this question below.

Question (*d*) is a matter of practicality. The test should be capable of being administered to the group without special equipment and personnel when these are not available. Here, economy of effort and material will be a factor to be considered.

Question (*e*) relates to the question whether the scores are stable and dependable; this is the reliability criterion. When the scores represent a comparable degree of control of the language in actual situations, it can be called *scoring validity* or, as Thorndike and Hagen call it, the *validity* as *signifying* or *construct validity*[2]. Below we will come back to this and the ethnolinguistic factors involved.

Question (*f*) is a question that we also have to ask seriously every time. The results of a test should be studied for information that might be of use in curriculum and/or syllabus revision, and in sharpening our techniques and methods of instruction.

2. **Validity in language testing**

The most common definition of validity one reads is: "the degree to which a test measures what it purposes to measure."[3] Another more practical definition is: "the degree of suitability and the usefulness of a test for our purpose." If we want to determine whether a test is valid or not we can try it out on a comparable group of persons and study the results in the light of the questions mentioned above.

These definitions bring up the questions as to what we want to test, how, and for what. The answers to these questions result in a recognition of several types of validity:

(1) *content validity* (also called rational or logical validity)[4]; we need to know what things we want to test and whether the test items are an adequate (and random) sampling of the "population," i.e. the components of the language as a whole, or a particular section of it.

(2) *predictive validity*[5]; we will want to know how closely the results of the test correlate with later accomplishment of the persons tested; for this we can compute the so-called correlation coefficient against a "criterion" measure, like the results of a group of persons of known ability who have taken the test before or are taking the test concurrently.

(3) *performance validity*[6]; we indicate herewith whether the techniques and the situations used really test the items that we want to test (cf. the above example on the distinction of / č / and / š /).

(4) *empirical validity*[7]; will the ability demonstrated at the test also be demonstrated in actual situations? Will the amount of control of the language be matched with the same amount of control in ordinary linguistic communication situations? More will be said about this below, because this is an important question which is frequently not given enough weight in language teaching and testing.

The first three types of validity can also be established by comparing the results obtained with results of the same test administered to a criterion group, e.g. native speakers of the language who by definition possess ability in the language. The fourth type can best be ascertained in an oral test in which the participants can be guided to demonstrate ability in the use of the items tested in a (near-) ordinary communication situation. Below we will suggest ways in which this can be done with an objective type of test.

The term *ability in a language* (or, *control of a language*) has been used several times above. We will define it here as the ability to use the appropriate elements and structures spontaneously and deliberately in actual communication situations. What we test is the control of a composite set of linguistic skills, in other words, proficiency in a language, in our case, in English as a foreign or second language. The population of our test items consists of the problems in the learning of English for speakers of a particular language and sociocultural background. We accept the definition of language learning as the acquisition of new habits, but include in the term habits such things as vocabulary (semantic features), paralinguistic features, and ethnolinguistic elements, especially when we also envisage some oral command as our teaching aim.

For the purpose of test construction and evaluating the validity of tests the problems are customarily divided into pronunciation, structural patterns, vocabulary, reading comprehension, auditory comprehension, etc. Most of the tests known to me (in Indonesia) are pruned of all cultural factors which are considered to be extraneous elements and to fall outside the scope of language.[8] This is partly justified by saying that we do not want to give an advantage to persons of a certain sociocultural background. However, if an overall command of the language is the purpose of teaching, we should rather include these things, because language does not exist in a vacuum but in a particular culture of which it is a part and a reflection[9]; these factors also include paralinguistic and kinesic features. Of course we need not test all the aspects and elements of a language in one test. In order to control the factors more easily we can allocate separate tests, or separate sections in the same paper, to the different aspects, like structure, auditory comprehension, etc., but among them there ought to be one paper or section on ethnolinguistic features in which we evaluate the person's control of the associations between linguistic expression and cultural and situational settings.

It is possible and meaningful to talk about the validity of the whole test, but it is also possible and more meaningful to talk about the validity of parts or items of a test, because a test consists of parts or, in the case of objective tests, of items. The validity of a test can be expressed, perhaps more meaningfully, in terms of the sum total of the validity measures of the component items. The questions concerning the validity of a whole test can also be asked of each item, e.g. "is the item included in the population" (content validity, etc.) In addition to those questions there are two other necessary and pertinent questions; the first is: does the item possess sufficient discriminating power? This is the most important in most types of tests; e.g. selection tests. The second is: is the item within the proper range of difficulty level? To answer these questions we need to make an item analysis, by which unacceptable items can be eliminated, and thus raise the validity of the whole test. An item analysis is within the grasp of the ordinary language teacher, because he does not need to have a special training to do this.

Discriminating power of an item is the degree to which it discriminates the better students from the poorer ones. The more common way of doing this is by taking the lower scoring 27 per cent (L) and the higher 27 per cent (H) of the participants. The number of mistakes in a particular item made by the lower group (W_L) substracted from the number of mistakes for that item made by the higher group (W_H) yields a number (D), in symbols $D = W_L - W_H$. Within the same number of options in one item, the higher the number D, the greater the discriminating value of the item. If $W_L - W_H = 0$, the discriminating value of the item is zero; if $W_L - W_H =$ a positive number, the discrimination value is said to be positive; and if $W_L - W_H =$ a negative number, the discrimination value is negative. Items with negative and zero discrimination values are eliminated. In some cases, zero discrimination may be meaningful in an ordinary classroom achievement test, where it may indicate that the problem has been equally well learnt or not learnt by everybody. The lowest discrimination value acceptable can be computed by means of a formula usually given in books on measurement and evaluation. Stanley[10] gives the minimum acceptable discriminating values in the form of numbers in a table which can be very easily used by the ordinary, statistically untrained, teacher.

If it is desired to obtain a still higher degree of validity, we can check the options or choices in each item for validity. For an option to be valid it has to attract (or perhaps, distract) a few students.

For a rough measure of the difficulty level of an item we have to compute the per cent of the number of mistakes made on the item with relation to the total number of candidates. A per cent which is not too far (±25) from 50 per cent is generally considered an acceptable level of difficulty. A combination of this with the more important discrimination value will give a measure of the suitability (acceptability or validity) of an item.

In our experience in Malang with university students, we have had to discard many items in the preliminary tryout of a test on the basis of an item analysis. In many of the discards we found other factors complicating the items over and above the features being tested. In many of these, it was found that cultural factors contained in the items were the chief culprits. As we will see below, we regard cultural features as integral parts of language; however, it will be necessary to test them separately in order that they do not invalidate the items constructed to test other features. For example, except in higher levels of proficiency, items on things concerning institutions peculiarly American like (American) football, the subway, foods, family problems, dating system, will pose complications.

This raises a problem concerning the content of our test; Should cultural factors be excluded from language tests?[11] If they are excluded from certain types of tests, are they also to be excluded from the whole battery in a proficiency test, a test that purports to measure the overall ability in a language? It is submitted here that we cannot, and should not, exclude them in an overall proficiency test. The next question is concerned with the methods of doing things. The answer that is closest to hand is an interview-type oral test in which we can present the candidate with the appropriate situations, both real and vicariously in the form of descriptions in the native language if necessary, in the elicitation of responses that we want to test. We all realize that this is a laborious and time-consuming method though more nearly like an ordinary language situation, for which the ordinary teacher cannot afford the time and energy. We can be a little more specific and ask the question of a proficiency test on English as a Second Language; what things do we need to test? What skills do we expect in a person possessing a good level of control of English? It is readily accepted that that person must have an acceptable, intelligible pronunciation, a comfortable ability in manipulating, both in writing and orally, the structure of the language (namely the morphemes, words, the arrangements of these), and the grammatical categories and linkages. He should be able to speak while keeping his attention on the content of the message. Although these statements contain only vague qualifications, let us accept them for what they are worth. What is more important at this point is the observation that we have left out completely any mention of those features that accompany speech, i.e. paralinguistic and kinesic features, and any reference to the socio-cultural setting or the ethnolinguistic features. Can we really say that we understand a language, if we fail to react appropriately to these features? And, should we not expect a certain amount of active control of these features of somebody who we say has control of the language? It seems to me that the answer to these questions is definitely affirmative. The use of a language does not take place *in vacuo*, and our speech is always accompanied and supplemented by body movements and modulations of voice and sound combinations which we do not customarily include in our linguistic descriptions.[12] Even in ordinary writing, which linguists often call an inexact representation of speech, we find indications of these in the form of special graphic symbols and verbal descriptions like using a double exclamation mark, underlining, capitals, and expressions like "he shrugged his shoulders," "he scratched his head," "hm, hm," etc. How this can be tested effectively is not very clear. Some suggestions are given below.

Imagine a play or novel without any of these paralinguistic and kinesic indications; I am sure it would sound as dry and lifeless as any lecture (including the present one), it would also lose much of its meaning. It seems to me that it is precisely because ethnolinguistic features are an integral part of language that the strict translation method has come to be regarded as an inefficient method. For example, when two English speaking persons meet they may say "How are you?" of which the Indonesian equivalent is "Apa kabar?" translated literally "What's the news?" to which a non-Indonesian might reply: "What news?" or "The U.S. has sent troops into Cambodia," etc.

When two persons are introduced it is customary in Indonesia to ask for what is regarded as personal (thus, by European standards, "private") information, like "How many children do you have?" If a person immediately talks about the weather without exchanging a few personal pleasantries, it will be regarded as strange or as an indication of lack of interest and may cause a feeling of discomfort in the other person.

3. Vocabulary and ethnolinguistic features

Before we conclude this discussion, let us make a brief digression to look at what language is in the context of language teaching and learning and hence also in the context of language testing.

De Saussure, by reducing language to its elements, namely signs, looks at language as a two-sided psychological entity represented by the "concept-sound image" diagram.[13] He describes the role of language as "a link between thought and sound."[14] He compares language to a sheet of paper with one side representing sound and the other side representing thought, neither of which can be divided from the other. He says further that this combination of thought and sound produces a form not a substance.[15]

We will use a modified form of this for our description of language (langue), namely as a system of links between meanings and sounds. In graphic form we can represent it thus:

meaning (translations)

---

language

---

sounds (writing)

In speaking we go from meaning to sound and in listening we go from sounds to meaning.

It is possible to study the sounds without relation to the other two factors. This is done in acoustic laboratories. The study of meanings as separate from language and sounds is done by social anthropologists. The study of language as a distinct entity from meanings and sounds is done in linguistics.[16] We as language teachers are interested in all three factors; in an authentic language situation we have all three factors involved simultaneously. In addition we have to add writing as another means of representation, and sometimes the use of equivalents in another language (indicated by "translation" in the diagram) if we use translation as an aid in, or method of, teaching. In the latter case we have the following diagram:

meaning

---

translation

---

language

---

writing

Here there is an additional stratum separating *language* from *meanings*, and the representation stratum is writing. However, the important thing that I want to stress is that the language teacher is concerned with *all three levels at the same time*, and that our aim is to make the students acquire control of the three aspects of language: meanings, linking system, and the representation system (either sounds or writing or both). If we at one time or another give more emphasis to one particular aspect, that is legitimate from a methodological point of view. However we should not do so at the expense of breaking the natural association between those aspects.

This is a danger that we have to guard against in drills in pronunciation and structure. When the association between language and meanings is broken, then we have a parrot-like situation. At this present stage of electronic technology it is not inconceivable that people will be able to construct a "talking" machine if they want to; this type of machine will react to stimuli (or cues) by a built-in system of "responses". This type of thing is what we often find as a result of a slavish adherence to the mechanical parts of the audiolingual approach. But this is a far cry from the situation of a speaking person.

Language consists of two levels of formal units: phonemes and lexical units (or vocabulary, i.e. morphemes, words, idiomatic expressions, intonations), and of patterns of relations (arrangements, linkages, categories). The phonemic system provides the link with the representation level (i.e. sounds or writing), and the vocabulary or lexicon provides the link with meanings, thus with the culture. On the whole, much more emphasis has been given lately in language teaching to the patterns of relations of structure and pronunciation, in particular in Indonesia; very often at the expense of vocabulary.

This emphasis on structure and pronunciation in language teaching is naturally reflected in language testing. Sophisticated tests on structures and auditory discrimination have been developed. The low degree of importance of vocabulary in teaching is reflected in the relatively unimportant place given to it in testing.[17] Through vocabulary we can test the control of the linguistic aspects of cultural behaviour or what we have called above ethnolinguistic factors, and understanding of paralinguistic and kinesic features. Some of these can conceivably be dealt with in a comprehension test, both auditory and reading. A proficiency test should include a part on vocabulary in which those factors are also dealt with.

We have seen that one of the most important questions in validity is "what things we want to test." The answer to this will determine the content of the test as a whole and of the items individually. When we can answer this question satisfactorily we will have a criterion for evaluating the content validity of the test battery, of the parts, and our items in each part of the test.

If control of the features mentioned above is a component of the control of a language (which we believe it is), then we need to test this in a proficiency test. Modern contrastive analysis has produced fairly accurate and complete lists of the linguistic problems in foreign language learning, which constitute the general population of a foreign language test. When we extend the contrastive analysis to paralinguistic and ethnolinguistic features, we will learn to distinguish the emic and etic units in the target language (English) and the native language. Here, we must admit that not much has been done in this field,[18] however, we need not wait for perfect and complete results before we start using the results of research. The little that we know about those features needs to be presented in teaching and therefore tested in evaluation. An example of how paralinguistic features can be tested is found in Lado[19], in which an understanding of *whistling loudly* in the situation of a soccer match in Spain is tested. Likewise, we can test an understanding of *sighing, laughing, saying something under one's breath, whispering, gasping, etc.*

An understanding of cultural factors (i.e. behavior and situations) can also be tested. For examples see Lado pp. 286-288. What is more elementary, hence more important, are the ethnolinguistic features, namely the things said peculiarly in a language by the constraints of the culture. The question arises as to how this can be done. As mentioned earlier, an oral test is the most immediate answer to hand. If an oral command is part of the teaching objectives, then an oral test is very necessary to have a truer evaluation of achievement and proficiency. Oral tests are time-consuming and difficult to score. Therefore ways should be found for doing the job with paper-and-pencil techniques which allow objective scoring. By way of examples, a few are given below. They are intended to cover a broad sweep of paralinguistic and ethnolinguistic features. The sentences in square brackets can be given in the student's native language if their level of proficiency is not so high that misunderstanding might not occur.

(1) He did not know the answer, so when he was asked he
{ shook / shrugged / raised / lowered } his shoulders ___

(2) [When he saw the state his table was in, he said:]
"That's very funny." [That indicates that he saw
something { he likes / humorous / interesting / unusual } there.]

Examples (1) and (2) look like an ordinary vocabulary test. In example (1) we have not only what is usually called "idiom" in language teaching but also a kinesic phenomenon. Example (2) is an instance of the different semantic areas covered by words which are usually regarded as equivalents in the two languages. Another example which happens to be authentic, on the word "funny" is the following:

(3) [Some Indonesian women went to visit an American friend to see her new-born baby-boy. One of them said:]

"His face is funny." [The mother kept silent, because she understood her as saying that

1. the child's face is humorous
2. the child is smiling
3. the child is pretty
4. the child looks strange

(4) [When she saw her friend wearing a pretty dress, she said:]

"Your dress is very beautiful." The friend answered:

"That's true"
"It's not worth mentioning"
"Thank you"
"It's not expensive"

(5) A: He doesn't like bananas, (does he)?

B replied, meaning to agree with A:

Yes, he doesn't
No, he doesn't
No, he isn't
Yes, he does

4 Conclusion

As we said above, speaking occurs in a setting or context, both linguistic and nonlinguistic. It is accompanied by paralinguistic and kinesic elements. All of these factors are culture-bound; in different cultures we have different systems of language, of sociocultural constraints, paralinguistic and kinesic systems. A person who has control of a language possesses control of these things. Therefore when we want to test control of language we will also need to test control of these factors, if our test, or battery of tests, is to be valid.

Tests can be partial and more specific. We can construct a language test into a battery consisting of language structure, pronunciation, reading comprehension, and aural comprehension, spelling, and vocabulary dealing also with the ethnolinguistic system, even conceivably the paralinguistic system. If necessary, on higher levels of proficiency, ethnolinguistic elements and paralinguistic features can be dealt with in a separate test part. Only with the inclusion of these features can a proficiency test be considered valid.

Footnotes

1. cf. R. Lado: *Language Testing*, 1961, p. 57.
2. R. L. Thorndike and E. Hagen: *Measurement and Evaluation in Psychology and Education*, 1961, pp. 161 ff.
3. cf. R. Lado: *op. cit.*, p. 321; D. W. Grieve: *English Language Examining*, 1964; pp. 23 ff.
4. Thorndike and Hagen: *op. cit.;* p. 161.
5. J. C. Stanley: *Measurement in Today's Schools*, 4th ed., 1964; p. 160.
6. Grieve, *op. cit.;* p. 23; cf. also Lado: *op. cit.;* p. 324.
7. cf. Lado, *op. cit.;* p. 324.
8. This refers in particular to the objective tests used in the English Department, IKIP Malang.
9. Especially in its vocabulary system. cf. E. Sapir: *Culture, Language, and Personality*, 1962; p. 36; and H. J. Mueller: "Cultural Reflection in German," in *Report of the Eleventh Round Table Meeting on Linguistics and Language Studies*, B. Choseed, (ed.), no. 13, 1962; p. 142; cf. also R. Lado: *Language Teaching*, 1964; pp. 23, 56.
10. Stanley, *op. cit.*
11. Here, it does not help to use the scores of native speakers as a criterion because they are either not bothered by such factors or they come from a certain (but different) subculture.
12. These, I take it, are in the border area between the things talked about in linguistic descriptions and the elements of cross-cultural understanding mentioned in Howard L. Nostrand's "Describing and Teaching the Sociocultural Context of a Foreign Language and Literature," in A. Valdman (ed.): *Trends in Language Teaching;* 1966.
13. F. de Saussure: *Course in General Linguistics*, translated by Wade Baskin, 1959; pp. 65, 114.
14. *ibid.* p. 112.
15. *ibid* p. 113.
16. cf. Morris's "Syntactics" in Charles Morris: *Signs, Language, and Behavior*, 1946.
17. The importance of vocabulary expansion at an advanced level is recognized by Lado in *Language Teaching*, pp. 129–30. It seems to me that postponing vocabulary building too long in foreign language teaching will affect student's interest adversely.
18. See, however, the bibliography of George L. Trager: "Paralanguage: First Approximation," in Dell Hymes (ed.) *Language in Culture and Society*, 1964; pp. 280–288.
19. Lado: *Language Testing;* p. 286.

TESTING READING SKILLS

by B. DEANE SMITH

The main purpose of this paper is to review some of the techniques used in the testing of reading, particularly at advanced levels, and to discuss the suitability of particular techniques for specific groups of learners of English as a second or foreign language. Before considering the testing of reading, however, it is necessary to identify at least the main skills involved in the reading process.

The Reading Process

One must begin, of course, with the assumption that the learner is familiar with the English graphic system; i.e. its alphabet, its punctuation, and its use of sentences and paragraphs.

The basic linguistic skills involved in reading English are simply stated: they amount to understanding the vocabulary (i.e. lexical items) and grammar (i.e. morphology and syntax) of the written language. Although easy to state, however, they are not so easy to achieve. It should be noted that "understanding grammar" is included as a basic reading skill. It is not sufficient merely to "know the words", because in written English words do not occur in isolation except in vocabulary lists or dictionaries. They occur in longer units, variously termed phrases, clauses, and sentences; and their form and arrangement within these longer units constitute the grammar of the language. Meaning is conveyed by the total context, both lexical and grammatical. At advanced levels in particular, reading comprehension is dependent to a considerable extent on the understanding of the complex syntactic patterns characteristic of formal written English.

Other skills are also needed by the advanced reader of English. These include the ability to identify the main theme or argument of a piece of writing, and to recognise ideas which are subordinate to the main theme, or examples which serve merely to illustrate it. If the writer's purpose is to provide factual information (as in a textbook, or article in a professional journal) the ability to draw correct conclusions from the material is important. If, on the other hand, the writer is putting forward a point of view on a controversial topic, it is useful to be able to separate fact from opinion, and to draw one's own conclusions. Clearly these abilities go well beyond the basic linguistic skills, and depend to an appreciable extent on both the general intelligence and previous training of the reader. If the learner of English as a foreign language has not received training in these advanced reading skills in his own language, he will find it more difficult to apply them in reading English.

The reader who specialises in English studies at the university level, and who has to deal with imaginative writing of different kinds, must be able to identify the author's attitude towards his subject, which sets the tone of the writing; and also the stylistic devices used by the author to convey his ideas. Skills of this nature, however, are not normally required of readers who use English as a medium of instruction in technical or scientific subjects.

It can be seen that the reading process comprises a complex of many skills. In practice a good reader employs most or all of these skills simultaneously, because they are mutually dependent, and have been separated here simply for the purpose of analysis.

Lastly, fluency in reading must be mentioned. Fluency (by which is meant reading at a good speed with full comprehension) can only be attained when the basic linguistic skills at any particular level have been mastered, and comes with extensive reading practice at that level. A high degree of fluency in reading advanced material is very difficult for a learner of English as a foreign language to achieve. In fact, reading fluency is an ability requiring training and practice even in the case of readers whose mother tongue is English.

What to Test

Reading has been defined as "understanding the whole sense of a sequence of written or printed words". The operative words in this definition are "whole sense" and "sequence", and it is clear that the definition is intended to cover the reading process as a whole; i.e. as a combination of skills. It follows that questions designed to test the overall reading process, the traditional "comprehension questions", likewise test a combination of reading skills, rather than any one skill in isolation. This is true of tests even at an elementary level, where the correct response to a question on the meaning of one short and simple sentence requires an understanding of both the lexical and grammatical context.

But there are testing situations in which general comprehension questions can be usefully supplemented by more specific questions, testing either vocabulary, or grammar, or both.

It is often convenient to test individual vocabulary items occurring in a test passage by isolating them in some way, and testing the reader's knowledge of their meaning *in the context of the passage,* usually by some multiple-choice technique. The disadvantage of this method of testing vocabulary is that one is limited in the choice of items to test. On the other hand, if the passage itself is well chosen, presumably the vocabulary items will be relevant to the needs and attainment level of the candidates for whom the test is intended. The advantage of vocabulary testing of this type, at least in tests of reading proficiency, is the high degree of contextualization of the items, which in turn increases the reliability of the test.

In reading tests designed for native speakers of English, knowledge of grammar tends to be taken for granted. In tests designed for learners of English as a foreign language, however, it is frequently desirable to include questions on the understanding of grammatical features which, for various reasons, are a source of difficulty for the learners concerned. The complexities of the verb system in English cause serious comprehension problems for many Asian learners, for example. In advanced reading material of a scientific or technical nature, noun phrase structures are another source of difficulty. Learners who have spent many years reading simplified material are inclined to lose their way in the long and involved sentence structures which they meet in their unsimplified university textbooks. One way of testing the understanding of complex sentence structures is by asking questions on pronoun references. These are merely a few representative examples of general grammatical areas which should be tested, if they lead to problems in reading comprehension for particular groups of learners.

How to Test

In selecting a passage for testing purposes, decisions regarding the length and content of the passage will depend on the level of the candidates, their particular English language requirements, and the length of time envisaged for the test administration. In general terms, a passage should contain at least 100 to 150 words, in order to provide a reasonable number of test items. At an advanced level much longer passages may be used, if time permits. In reading proficiency tests it is normal to use several passages, varying in length, content, difficulty and style. Each passage chosen should be self-contained, i.e. meaningful in itself, and the questions set on it should be capable of being answered through intelligent reading of the passage alone, without the use of any outside knowledge. It follows that passages of a technical or otherwise specialised nature should be avoided, unless of course the test population itself is a specialised group.

Questions may be placed in one of two broad categories, depending on the nature of the response required of the reader.

A. **Production**

In the category which can be called productive, the reader is required to produce an answer to a question, either orally or in writing. Production may be free, or controlled to a greater or lesser extent.

Example: Beginnings of Government in Europe[1]

About 4000 years ago, the Greeks began to move across from their homes in the mountains of Asia towards the Western Mediterranean Sea. At this time they were *wandering* shepherds, living in groups of families called tribes. They had no real government: such a thing was unnecessary because there was no *public* business, there were no *taxes*, and nobody *owned* any land, since the tribe moved from place to place the *whole* time looking for grass for its sheep and goats. There were also no law cases and no lawyers, and men were controlled by a few customs, such as the one that said that if one man killed another, one of the dead man's *relations* had to find and kill the killer.

1. Give for each of the following words another word or phrase of similar meaning to that in the context.

   taxes owned relations

2. Write down the following words and then give for each a word or phrase of *opposite* meaning to that used in the passage.

   wandering public whole

3. Give brief answers to the following questions using ONE COMPLETE sentence for each. Use your own words as far as possible, but your answer must be based only on the information in the piece.
   - (*a*) Why was a government unnecessary to the wandering Greek shepherds?
   - (*b*) What example of an early custom does the writer give?

---

[1] Slightly adapted from Hill and Fielden *Further Comprehension and Precis Pieces for Overseas Students*, pp. 53, 55.

Comments

1. If recognition is to be tested through production, the basic assumption must be that the learner's productive skills are at least somewhere near his reading skills in level of overall proficiency. This is normally the case with native speakers of English. It is often the case with learners in the so-called "TESL" countries, where English is taught as a second language; but only occasionally is it true in "TEFL" countries, where English is taught as a foreign language, and where, in general, the speaking and writing skills tend to lag behind the reading skills to a considerable degree. Although the vocabulary questions in the example (Nos. 1 and 2) might work reasonably well in such countries, an invitation to "use your own words as far as possible", as in question No. 3, is likely to result in the learner's using his own version of English grammar as well, which may lead to more harm than good.

2. A problem thus arises in the evaluation of answers in productive-type texts. If meaning is conveyed by both lexical and grammatical context, a decision must be made as to whether grammatical errors in a candidate's answer should be permitted to affect his score. Opinion on this point is by no means unanimous, and this introduction of a subjective element in evaluation means that productive-type items are unsuitable for test administrations where large numbers are involved, and the grading is carried out by more than one person. Furthermore, the grading of such items is very time-consuming.

3. The foregoing comments should not be interpreted as a blanket condemnation of productive-type comprehension tests, even in TEFL countries. They can be used successfully in a classroom testing situation, where numbers are small, and both the teacher and his students understand clearly what is required with respect to the nature and length of the response, and the method of assessment. At very elementary levels, for example, a short response such as "Yes, it is", or "No, it isn't" may be all that is required. Questions may be framed in such a way that the answers can be lifted almost word for word from the text of the passage, with or without the help of guide words. But such procedures are teaching methods rather than testing techniques, if only because a teacher will normally follow up a test in a succeeding lesson, thereby using the test as a teaching aid. For the reasons already stated it is difficult to justify the use of productive-type reading test items in any situation where large numbers of candidates are involved; and in countries where English is taught as a foreign language, the degree of imbalance which so often exists between productive and receptive skills makes the testing of reading through writing largely irrelevant.

B. Recognition

Receptive-type reading test items can be placed in this category, in which the reader is required to recognise the correct answer among several given, to state whether a given answer or statement is true or false, to re-arrange a number of statements in their correct order, to match words with definitions, and so on. In questions of this type a *choice* is either explicit (as in the typical multiple-choice item) or implied. The greatest single advantage of the multiple-choice item is that the correct answer is fixed by the test-writer, and the opinion of the marker of the test does not enter into evaluation. This remains true, whether the correct choice is in fact the only right answer, or whether it represents the best or most complete answer among those given. Thus it is possible for test papers to be graded rapidly, accurately and impartially, sometimes by mechanical means.

It is not the purpose of this paper to defend the use of multiple-choice items in the testing of reading. They are already well established as a testing technique, and their use is essential in many contemporary testing situations, where test populations may run into the thousands. It is true, however, that a heavy burden falls on the test-writer in the construction of multiple-choice items, and valid criticisms may sometimes be made of items which, for one reason or another, are unsatisfactory. Any testing technique, like any teaching technique, can be abused as well as used.

A common criticism of multiple-choice items is that they are "too easy". In fact, items can be made difficult enough to test even a mature native speaker of English; but it is possible for poor construction to turn an otherwise suitable question into a test item which is too easy for a particular test population. Observance of the following general rules of item construction will help to avoid this danger.

1. No reading comprehension item should be able to be answered correctly without careful reading of the test passage itself. This means that all the choices in any one item should appear reasonable within the context of the item.
2. All the choices offered in an item, whether correct or incorrect, should attract at least some candidates, if only the weakest. It sometimes happens that one of the choices fails to attract any candidates. This may result from the fact that the choice is obviously wrong, or absurd. In the typical test item offering one correct choice in a total of four choices, this has the effect of increasing the candidate's chances of guessing the correct answer from one in four (i.e. 25 per cent) to one in three (i.e. 33 per cent). If two choices fail to attract any candidates, the item is reduced to the true/false level, where there is a 50 per cent chance of guessing the correct answer. Such an item would be statistically unacceptable, of course. The incorrect choices in an item are called "distractors" or "decoys"; and the strength or weakness of an item as a testing device depends to a large extent on whether the distractors fulfil their function. It is seldom possible to determine the efficiency of distractors merely by inspection; pre-testing of items is essential, so that non-functioning distractors can be eliminated.
3. If an item can be answered correctly merely by matching words in the passage with words in the correct choice, this is clearly too superficial a procedure even to be called a test of reading comprehension, and should be avoided. Even when the meaning of only a single phrase or sentence is being tested, a synonym or paraphrase should be used in the test item.

The above general principles concern the writing of test items. The meaning of scores on multiple-choice type tests should also be clearly understood. One weakness of these tests has already been mentioned, namely the element of guesswork which can enter into the candidate's performance. This makes questions of the true-false type unsuitable for reading proficiency test, although they can be useful as a teaching aid in a classroom situation.

In an experiment to determine the scores actually made by sheer guess-work on a 100-item multiple-choice test, with four choices per item, the answer sheets for such a test were distributed to a large group of persons, whose instructions were to score the sheets entirely at random, without reference to the actual test materials. Inevitably the mean score was 25 per cent. However, the range was from

18 per cent to 36 per cent, and the conclusion reached was that any score up to about 40 per cent was very likely to be unreliable as a measure of the skills tested. Indeed, scores in excess of 50 per cent are really necessary for an acceptable degree of reliability. This, coupled with the fact that it is theoretically possible to score 100 per cent on an objective test, means that in cases where such tests are used for making important decisions, the "passing" grade is normally set quite high, in the region of 70 per cent to 80 per cent.

Examples of Multiple-choice Test Items Used at an Advanced Level of Reading Comprehension

The reading material used in the examples in this section is taken from R. A. Close: *The English We Use For Science* (Longmans), although the test items are the writer's own. The items have been placed in three categories: General Comprehension, Vocabulary and Grammar. The categories are not mutually exclusive. Naturally any question testing the understanding of the main ideas or concepts of a passage involves both vocabulary and grammar. Furthermore, the distinction between items testing vocabulary and grammar is not clear in every case, and sometimes amounts to no more than a difference in emphasis, or focus of attention. The categories have been adopted in the present instance for convenience of presentation.

A. General Comprehension

Example:

I intend to deal with the problems of formulating practical and realistic plans for the application of science and technology by the government of a newly emerging country. I will specifically have in mind the new African nations, since for various historical reasons, they are particularly short of trained manpower.

The writer begins by stating that he will

(*a*) deal with the historical reasons for the emergence of the new African nations.
(*b*) discuss scientific and technological planning in developing countries.
(*c*) discuss the shortage of trained manpower in the new African nation in particular.
(*d*) formulate practical and realistic plans for the application of science and technology.

Comments

1. This is an example of the "sentence-completion" type of item. The first part of the sentence, known as the "stem" or "lead", should provide sufficient context to clarify the nature of the problem. From this it follows that careful reading of the "lead" should be necessary for the correct completion of the sentence, in addition of course to the understanding of the passage. The key word in the lead is "begins" so the examinee should direct his attention to the information given in the opening paragraph, which is here quoted in full.

2. The correct choice, (*b*), is a summary of the information contained in the first sentence of the paragraph. The use of the phrasal verb "deal with", which occurs in the passage, might have provided a too obvious clue; so it is replaced by "discuss".

3. On the other hand the incorrect choice, or distractors, do make free use of the actual words and phrases occuring in the passage, as in this way those

candidates who tend merely to match words and phrases, without fully understanding the significance of their use by the author, may be tempted to make the wrong choice. Distractor (*a*) for example does begin with the verb "deal with", although it would not be chosen by a reader who understood the writer's sequence of ideas, and the relationship between them. Distractor (*c*) is incorrect because "the shortage of trained manpower" is merely the reason for the author's focus on the new African nation; it is not in itself the main theme of the opening paragraph, nor indeed of the passage as a whole. Distractor (*d*) actually uses eleven words lifted directly from the passage in their original form and order, and is specifically designed to attract the superficial reader. The careful reader, however, will avoid it because he has understood that the writer is only discussing the problems of formulating plans, and is not necessarily a planner himself.

4. It should be noted that all the choices, both correct and incorrect, are feasible within the context of the item. In the first place, they are all grammatically acceptable continuations of the first half of the sentence. This point, which might be considered too obvious to be worth mentioning, is in practice sometimes forgotten by inexperienced item-writers. In the second place, not one of the choices is obviously absurd or irrelevant, to be eliminated without further reading of the passage. No distractor is totally unrelated to the passage; the meaning content of all the distractors is drawn from the material of the passage itself, a natural consequence of using a large proportion of the actual vocabulary items used by the author.

5. Another consequence of using the author's vocabulary to a large extent in the writing of the item is the fact that the general level of difficulty of the language used in the item is the same as that of the passage. In fact, the language of the item can be made simpler, if desired; but as a general rule it should not be made more difficult, as irrelevant comprehension problems may then be introduced in the understanding of the choices offered in the item itself.

6. A final point worth noting is that, although the purpose of the item is ostensibly to test "general comprehension," a degree of specification has been introduced with the words "The writer begins by stating that . . . ." general comprehension questions should not be capable of being answered from the reader's general knowledge, without reference to the passage. Even in cases where the theme of the passage is of general interest, questions can be made more specific by requiring the reader to identify features of the author's own treatment of the theme, particularly his presentation of facts and ideas, the relative importance which he assigns to them, and the conclusions which he draws from them.

B. **Vocabulary**

Examples:

The small size of the average farm . . . would have been a serious obstacle to these changes in *farm practice;* but industrialisation, with its higher wages . . . is already attracting labour from the *rural* areas. A 35-40 per cent *fall in the farming population* by 1970 is *forecast.*

1. *obstacle*
   (*a*) threat
   (*b*) objection
   (*c*) obstruction
   (*d*) danger

2. *rural*
   (*a*) urban
   (*b*) suburban
   (*c*) city
   (*d*) country

3. *forecast*
   (*a*) promised
   (*b*) prophesied
   (*c*) calculated
   (*d*) required

4. *farm practice*
   (*a*) farming methods
   (*b*) practical farming
   (*c*) practice farming
   (*d*) farm produce

5. *fall in the farming population*
   (*a*) decrease in the number of farms
   (*b*) increase in the size of farms
   (*c*) decrease in the number of farmers
   (*d*) increase in the number of farmers

Comments

1. All the vocabulary items in the example worked well with a group of advanced learners of English as a foreign language, in at least one case, rather surprisingly well. Item 2, for example, is on the face of it a poor item, with its use of the synonyms "urban" and "city". The fact that they cannot both be right should in theory eliminate them at the outset. However, in practice the item proved an excellent discriminator, with all the distractors attracting at least some readers. Subsequent enquiry revealed the reason: three of the words in the item, viz. rural, urban, and suburban, were known to relatively few of the test group, and the success of the distractors was due more to blind guess-work than to anything else. Item 3 also proved to be a good discriminator, but for the wrong reason: the word used in the correct choice, "prophesied", was unknown to almost all the test group.

2. As can be seen, the term "vocabulary item" includes phrases as well as single lexical items. Test items 4 and 5 are examples of these. The use of phrases as test items often permits greater flexibility in the construction of distractors, as can be seen from Item 5.

3. As a basic principle, the vocabulary items selected for testing purposes should be important to the understanding of the writer's main flow of ideas.

4. Another important principle in the construction of vocabulary test items is that every choice offered in the item should fulfil the same grammatical function as the problem word or phrase. In other words, a noun should have noun distractors, a verb should have verb distractors, and so on. Further, nouns should correspond in number, and verbs in tense and voice. In this way no purely grammatical clues are offered in a vocabulary test item; the problem should be one of lexical meaning only.

5. When vocabulary is tested in context, as it is here, it tends to be a less "pure" test of the reader's knowledge than is a test of vocabulary items in isolation. This is because contextualization in itself frequently enables the reader to infer the meaning of a word or phrase from the general sense of the passage. However, it would probably be agreed that the ability to infer meaning from context is an important component of reading skill, and as such its incorporation as a testing technique is fully justified. Furthermore, it is an ability which experience indicates is not so widespread as teachers and testers of English as a foreign language sometimes assume.

6. In cases where a "word" has several meanings, depending on context, it will sometimes be possible to devise at least one or two distractors from other contextual meanings of the same word. The distractors will be likely to work,

however, only if the other contextual meanings are in some way related to the meaning the word has in the test passage. Totally unrelated meanings will not provide useful distractors for any except the weakest readers.

C. **Grammar**

Examples:

Of other creatures, *all* that we have found are the footprints that they have left in the wet mud . . .

1. *all* refers to

   (*a*) creatures (*c*) footprints
   (*b*) we (*d*) they

   . . . the tail (of the Archaeopteryx), instead of being as a modern tail, fixed to its pygostyle, or what is more popularly known as the "parson's nose", was long like that of a reptile and to *it* feathers were attached on either side.

2. *it* refers to

   (*a*) the tail fixed to its pygostyle (*c*) the "parson's nose"
   (*b*) a reptile (*d*) the tail of the Archaeopteryx.

Comments

1. Both these items are of the "pronoun reference" type, and between 40 and 50 per cent of a group of advanced learners failed to answer them correctly.

2. In Item 1, the pronoun precedes the noun to which it refers. In Item 2, it follows its noun at a very considerable distance. The correct understanding of pronoun references is a convenient check on the understanding of complex sentence structures.

Example:

The geothermal power station at Wairakei in New Zealand is *the first natural-steam-driven generating* plant in the southern hemisphere.

3. *The first natural-steam-driven generating plant*
   (*a*) the first plant for generating natural steam
   (*b*) the first natural plant for driving steam
   (*c*) the first generating plant driven by natural steam
   (*d*) the first steam plant for driving generators

Comment

Complex noun phrase structures of the type tested in Item 3 are not uncommon in scientific and technical writing. When they occur, they can often produce useful test items.

Other examples could be given of items testing important grammatical areas such as tenses, and different kinds of subordinate clauses. These are well-known problem areas, however, and should not require elaboration here. It is hoped that the few examples given may suggest other ways in which the understanding of grammatical structure may be usefully tested.

A RESEARCH PHILOSOPHY AND ENGLISH LANGUAGE TESTING

by PETER B. BOTSMAN

Among practising teachers and administrators, there is a tendency to regard research and philosophy as alien and "unspeakable rites"[1]. It would seem therefore that anyone foolish enough to title a paper using both of these terms is likely to be dismissed as a "boffin" or a "weirdo" — a dreamy idealist far removed from the pressing realities of day-to-day teaching in the classroom. I wish to begin by assuring you that I am still engaged in teaching the morphological mysteries of the English Language for thirty-five weeks in each year. However for the rest you must judge for yourselves.

But I make no apology for the title of my paper even though I recognise that too often researchers and even more philosophers are apt to be considered as somewhat akin to Joseph Conrad's "pilgrims" — hapless lost souls groping blindly in some impenetrably obscure "heart of darkness". Indeed many think them worse even than this for they are seen not merely as passive idolators, some of them are active, some have been known to solicit intelligent young people to join them in their esoteric rites.

Be warned then, lest you be beguiled. I charge you to adopt the role of Marlow in Conrad's novel. Be the practical man seeking purpose and clinging tenaciously to it, for I shall try to entice you to take at least a few steps into the lush jungle of theory in the time that I have to speak to you.

The nature of theory

There is in the minds of many people a basic misconception about theory and it is probably this misconception that prompts teachers to distrust both theory and those who practise theory building. In essence this misconception is aired in the sort of objection to a new idea or a new way of doing that is prefaced by the words: "That's all very well in theory, but in practice . . ." What is implicit here is the notion that theory and practice are polar extremes separated by an enormous gulf which can never be bridged. I want to stress that this dichotomy is often only imagined and that a continuance of this sort of misconception can only lead to mental stagnation and the *status quo* conformism that has characterised educational practice for far too long.

Toward a Science of Education

At this point let me recall for you the words of John Dewey. In his *Sources of a Science of Education,* he comments: "Theory is in the end . . . the most practical of all things, because the widening of the range of attention beyond merely purpose and desire eventually results in the creation of wider and farther reaching purposes and enables us to make use of a wider and deeper range of conditions and means than were expressed in the observation of primitive, practical purposes." Notice, by the way, Dewey's penultimate adjective. It may sound harsh to call practice 'primitive', but Dewey means it, I'm sure, quite literally. As its root suggests, 'primitive' means a first or basic step or stage. I want to persuade you that as teachers we can never be content with primitive practice, we must take further steps.

---

[1] For a detailed statement of this phenomenon and some pertinent comments on its endemic causes, see A. P. Colardarei and Jacob W. Getzels, *The Use of Theory in Educational Administration,* Stanford, Calif., Stanford School of Education, 1955, p. 10.

The first step beyond practice occurs when we start to think that perhaps our practice can be improved. Once we start to think about how a practice can be improved we have taken the first step towards theory for as I shall try to show very briefly one form of theorising may be no more nor less than a systematic way of thinking about ways of improving practice. If we accept this it follows that if we deny a place to theorising and theorists we are in effect condoning the preservation of the *status quo*, for carried to extremist lengths the denial of theory is tantamount to eliminating the possibility of any sort of progress or the discovery of any further meaning in anything, which leads us to the absurdist position so delightfully caricatured by Lewis Carroll, " . . . if there is no meaning (as the King said to Alice in Wonderland) that saves us a world of trouble, you know, as we needn't try to find any."

Requirements of Educational Theory

But I am in danger of finding myself in a 'world of trouble' unless I confine myself to my major task which is at this point to establish what we should expect of theory and its relationship with practice in education. To turn to the nature of educational theory, I suggest that theory is a purposive form of thinking that we employ when we are attempting to solve a problem, that is, it is not indulged for its own sake, it should always be an attempt to find answers to real problems. The second part to stress about theory is that it always represents a systematic form of thinking; that is to say, a systematic correspondence between the abstract symbols used in the theory and their referents in the real world is established and maintained. This close correspondence gives rise incidentally to the use of the word 'model' as a synonym for theory, as in Robert Dubin's thesis that the goal of all scientific activity is to "build viable models of the empirical world that can be comprehended by the human mind."[2] Assuming for our present purposes that scientific activity means systematic and purposive enquiry, it seems worth pursuing Dubin's idea a little further.

The first word that needs comment is the term 'viable'. As we are all aware this means literally 'capable of maintaining life', but as used in context here, where it is coupled with 'model', it implies that the model is connected with the real world in such a way that its vitality is assured. Put simply this means that a theory or model that is not intimately connected with the real world — the world of practice — is not under the term of this definition a viable theory, merely, one presumes, a form of intellective abstraction.

The second set of terms in Dubin's phrasing that we should consider further are the words, 'the empirical world' and the qualifying clause that accompanies them. We have noted already that the world as evidenced in our experience, is what a model seeks to parallel. Logically therefore the model must seek to contain as many of the identifiable operative factors in this world as is possible. But just what is possible in a model is suggested by the limiting clause, "that can be comprehended by the human mind" for this qualification simply allows that the world is infinitely more complex than the human intelligence that seeks to fathom its mysteries. In short, most models in order that they may be comprehensible to the limited human mind, are necessarily simplified.

---

[2] Robert Dubin, *Theory Building*, New York, The Free Press 1968, p. 9.

Let us now pause to review the special conditions that we should ask of theory in education. They are (*a*) a condition that ensures purpose, (*b*) a condition that ensures connection with the real world, (*c*) a condition that the model replicates as many of the identifiable factors impinging in the real world as possible and (*d*) a condition that requires the model to be sufficiently simplified as to render it susceptible to human comprehension. Now all these distinguishing characteristics are conditions that ensure that theory building is closely integrated with real problems in the real world. Hence it follows that if these conditions are met the alleged or implied dichotomy between theory and practice disappears. However it would be foolish and unrealistic to claim that all that poses as theory meets the conditions we have established. Moreover it is not hard to find reasons why many theories fail to recognise them as critical factors. But rather than attempt to justify failure to meet these conditions, the point I wish to stress is that if theory and theorists in the field of education are to be called to account, they should be judged by their ability to meet the conditions specified, not prejudged because of the existence of a widespread and widely believed prejudice.

A Changed Attitude to Theory

If theories and theorists were judged by these standards, I believe that there would be two immediately beneficial products. The theorists would be reminded of their primary responsibilities, forced to state their aims more clearly and generally made aware of the mutually dependent relationship of theory and practice, and on the other hand those who presently accept the crude notion that theory and practice are dichotomous states, might well be persuaded to take a more lively interest in both theory and the research projects that are normally allied to it. Come the millenium, you say, but I firmly believe that such a situation can and must be achieved.

I suggest to you that in teachers' colleges and in in-service training programs the men and women in our teaching services must be educated to think of theory in the way I have outlined — that is, as a positive and practical process. Moreover they must be educated to think of themselves as theorists, not merely as instructors. This is not the place nor the time to go into detail, but to return to Dewey's thesis, I believe that if we are to build a "science of education" we must educate our teachers in this way "for science and research are but the accumulated and still accumulating form of man's most carefully derived understanding involving appropriate problem-solving methods in gathering and interpreting evidence."[3] An understanding of theory is the first step towards this objective. However, it is clearly not enough that teachers think of themselves *only* as theorists. If they are also to be "scientists" they must be consciously guided by a research philosophy in their work and this implies certain other obligations as I shall try to show presently.

Educational Practice

But having talked at length about theory, I am still not yet nigh to my central theme, because I must now say a few things about practice in education. Practices in education, particularly those practices commonly employed in classrooms are often called "methods" and a system of practices hence becomes a "methodology". This of course also applies to "method" of testing. Now there is nothing wrong with this simple change of name so long as we remain aware of a latent danger. The danger that I refer to here is the tendency to regard a methodology as some sort of ultimate goal, an end point to be achieved beyond which we can go no further. What I am suggesting is that the terms "a proven method" should be regarded with healthy skepticism if not downright distrust, for these terms imply that it is extremely unlikely that further refinement or improvement of the method can take place. This is just not true and I would even go so far as to say that anyone who claims to have "perfected" a method has probably only ceased to learn.

---

[3] Carter V. Good and Douglas E. Scates, *Methods of Research: Educational, Psychological, Sociological*, New York, Appleton-Century Crofts, Inc., 1954, p. 8.

I want to suggest that originally all methods began with a theory. The theory has become a method after testing, modification and refining — a sometimes very long and slow process of metamorphosis. I think this notion is demonstrably true, but we tend to lose sight of this fact and regard methods not as things that have evolved but as fixed and unchangeable. Teachers should be made aware not only that methods have evolved from experience, but also that further empirical discovery and modification is always possible.

Hence a strict conformity to method for its own sake must be as strenuously resisted as the tendency to disregard and discountenance theory. However it is probably true that this represents an easier task than the task of combating the entrenched antipathy towards theory. This is because intelligent teachers invariably adapt and modify methods both in the light of their own empirical experience and because of their individual differences in personality. In other words I believe that most teachers approach methodology pragmatically and flexibly which is as it should be, for educational practices that are codified into rigourously detailed syllabuses and teaching instructions which are then slavishly applied in the classroom should be shunned like the plague. In such a system progress and development is inhibited, teachers are degraded into mere pedants and most important of all the students who are trained under such a system, while they may be suited to Huxley's "Brave New World", are certainly not educated to take their place in the world in which we live today.

An Immediate Need

Now I want to make it clear that I am not advocating a visionary far-off ideal — I think that we should be working towards the situation that I am about to propose right now. I concede that I do not know the Southeast Asian world and in particular its educational system very well, but I am nevertheless confident that you have in your schools a hard-working, intelligent group of dedicated language teachers and I therefore feel sure that with wise administration these people could be enlisted as enthusiastic laboratory workers and research assistants who would be sympathetic towards and actively support theory and research once its practical nature is explained to them. In other words what I have rather pretentiously titled as a "research philosophy" may, in many cases, be translated into behavioural terms as an interest in and an understanding of theory, with a willingness to assist in empirical research in the classroom. I believe that this willingness will be manifest because it has been my experience in many places that once the nature of research, and the particular aspect of it implicit in any project, is explained, teachers are overwhelmingly cooperative and they are, after all, just as interested as we are in increasing their understanding of what we should teach and how we should teach. Thus I see the main task as convincing teachers of their vital role as "scientists" and I believe this to be by no means an impossible task.

So now I have talked at length about theory, I have suggested that educational practices should be constantly re-examined and modified in the light of experience and I have suggested that teachers should be enlisted in the service of empirical research, but I have yet to say more than a word or two about English Language Teaching and Testing which is, after all, the subject of this conference. I shall therefore try to remedy this situation forthwith.

Language Teaching Today

In language teaching today there is among theorists a spirited and often fierce debate. As W. A. Bennett has noted "feelings run high among the advocates of various linguistic and psychological theories and the alarms and excursions deep inside the respective territories are many and fierce."[4] But in the midst of all this furore stands the language teacher struggling with the ever present reality of his pressing tasks. Small wonder that he becomes impatient with theorists for they are clearly little concerned with him. So he goes on his conscientious way ignoring the linguists and the psychologists whom he suspects will continue throwing their verbal thunderbolts about in the cloudland, (and I use the term advisedly), far above his head, no matter what he says or does.

The situation I have described exists in Australia and even more so, I regret to have to say, in New Guinea where I am presently working. I may be wrong in assuming that it probably also exists in Southeast Asia; I hope I am, but I think I am not. In any case the situation I have described is wrong and must be remedied. And it is wrong precisely because theory and practice have been quite divorced. As I have already suggested educational progress is extremely unlikely unless they are positively wedded and this means not only that they have exchanged solemn vows, but that they are joined in deed. Our problem, as in most cases involving divorce, is to re-establish harmony.

The Role of the Language Teacher

My central thesis which I have come to at long last is that teachers in general, but particularly language teachers because of the "divorce" situation I have outlined, must become vitally aware of their central role as empirical scientists, aware not only of the nature of theory, but capable of formulating theories, capable of testing them and of critically examining the data produced as a result. This, I must insist, is not really the radical demand it seems, for language teachers whether they realise it or not are in fact partly performing this role whenever even the simplest form of classroom test is given. The problem as I have said is one of awareness and the lack of a "research philosophy".

However, this sort of lack can be remedied. Let me cite a simple illustration of how a beginning can be made. I was once asked to take as a teaching adviser a group of young teachers all of whom were having difficulty in teaching English to adult European migrants. I began with the simplest form of testing of all. After teaching a particular point in a normal lesson, a set of exercises requiring the use of a particular syntactic structure was set. The teachers wrote the set on the board and in addition I asked them to rule two vertical columns to the right of the set, one marked with a cross and one with a tick. At the end of the lesson the teacher merely entered the number of students who had achieved the correct answers and those who had not in the appropriate column. This practice I asked them to persist with for a period of two weeks. I then left the teachers to get on with their work for that period. You can probably guess what happened. When next I called on the classrooms the ideas came almost as fast as the questions, and from the comments passed it was obvious that more had been learned about foreign language teaching and learning in that two week period than the whole of the previous semester. These young teachers were at first anxious to ask me what I thought (what my theory was). Why, for instance, did 32 students get example (a) right and only 8 fail on that item, while on example (b) which appeared to present the same syntactic problem, only 20 scored the correct response? I encouraged them to try to supply their own explanations and their response was, after sometimes quite a few leading questions, to come up with a "theory" which I then encouraged them to test. To cut a long story short, through simple empirical data analysis these teachers were encouraged to develop critical insights both into their own teaching and the students' learning.

---

[4] W. A. Bennett, *Aspects of Language and Language Teaching*, London, Cambridge University Press, 1968, P. 1.

The Need for Theory-Practice Interaction in English Language Testing

Now I stress that this was merely a beginning. However, an important principle is illustrated here. You will recall that I suggested that there were four basic conditions we should demand of theory in education. All of course are important, but perhaps the central characteristic of all these conditions is the one that demands constant interaction between the theoretical model and the real world. Now the "real world" in the context of E.F.L. work is the classroom and the pupils and teacher in that classroom. In the example I have given you, what each of these young teachers did was to analyse the results of their teaching practices through the medium of a simple test, and by so doing they were persuaded to re-examine their "methods" and modify them in the light of the data produced. In short, they were testing a model in the real world, then modifying and refining the model or theory as a result of their empirical findings. This seems so simple as to be almost trite, but it is a fundamental principle that is applied in all research projects.

Now I stress again, this is only a beginning, albeit an auspicious one. The point, however, is that once teachers are made aware of and interested in the potency of this kind of "feed-back" as a means of progress, then we are well on the way towards finding a way forward. I could go on to outline ways in which, from this sort of beginning, a research-based "philosophy" can be developed, but I have no doubt that you can supply the intermediate steps without any difficulty for, if we remember that we are aiming to effect a synthesis between theory and practice, we will not stray too far from our objective.

Thus I repeat, an immediate task is to alert our teachers to their vital role as empirical scientists who are engaged in testing and validating theoretical notions in the real world of the classroom. If we can accomplish this task, I believe that E.F.L. teaching is poised to take significant steps forward. This is because few other areas in the behavioural sciences can tap the potential source of information flow such as might exist in language education where every classroom could be thought of as a laboratory. Moreover, in language teaching we have now available enough theory to keep us busily engaged in empirical research for a long while to come. In addition, the achievement of the sort of attitudinal climate that I have outlined has a generative effect. This is because contrary to popular belief the sort of empirical research that may profitably be done by teachers in their classrooms is not merely passive. I want to conclude this paper by showing very briefly why and how empirical research of this kind can be extremely productive of new insights, new ideas and better, more scientific educational practice.

The Products of Empirical Research

In an article in the American Sociological Review in 1948, Robert Merton pointed out that empirical research in the behavioural sciences should not be regarded as merely passive testing. He suggested that it may perform at least four other important functions. It may initiate. It may reformulate. It may deflect and it may clarify. Let us consider each of these in turn in relation to E.F.L. testing.

A theoretical model is established and tested, but on analysing the data produced as a result of the testing one often finds an unanticipated result which may be quite inconsistent with the original theory. Suppose that a teacher gives a test and on analysis the results show something quite unexpected. Given a research philosophy the teacher's curiosity will be piqued by this fortuitous by-product and one may assume that he will explore further. In fact this unexpected outcome may lead to a completely new direction of enquiry. Put in its simplest terms, a teacher who keeps careful records associated with his teaching procedures may be persuaded to employ a totally new method in order to try to take account of the extraneous unanticipated factor that his first test isolated. At more sophisticated levels of testing the same sort of thing can easily happen. Hence empirical research may be said to initiate or induce completely new approaches.

Secondly, it will be obvious even from the single instance I have cited that a reformulation of ideas is likely to follow thoughtful and critical analysis of test data. This in turn leads to changes in procedures.

Thirdly, empirical research attitudes towards testing may have the effect of deflecting the mainstream of thought from an existing course, or, put another way, the focus of theoretical interest may shift quite radically, as a result of research findings. Some evidence that this is already happening may be traced through the literature relating to the use of language laboratories and I think it requires little imagination to see that if guided by classroom realities such effects may channel rather abstract notions into useful teaching procedures.

Finally the pressure exerted by empirical research on the investigator is important. I need not emphasise to a group such as this the importance of careful definition of the aims and concepts associated with a research project. Suffice it to say that the need for precise formulation of concepts and objectives and the careful definition of aims is a most urgent need in education generally and specifically in the field of English Language testing. A research attitude does exert considerable presure on the investigator to think about why he is doing what he is doing and in my view can only lead to a clearer understanding of our educational task.

Thus in all these ways empirical research goes far beyond mere passivity. I should like to conclude this paper by suggesting therefore that empirical observation and investigation in the classroom by teachers who are sensitised towards theory and willing participants in research projects should be recognised as an immediate and urgent need. This conference has clearly demonstrated that at E.F.L. seminars the greatest attention is perhaps inevitably given to the phonemic, lexical and grammatical aspects of foreign language learning. Doubtless this is because it is felt that it is in these areas that a breakdown of communication is most likely. I have tried to suggest that an even more fundamenal breakdown, that between theory and practice, is at present clearly apparent in the teaching and testing of foreign languages and to those who dismiss this as unimportant I can do no more than address these lines of the English poet, John Dryden,

"Error like straws upon the surface flow;
He who would search for pearls must dive below"

OBJECTIVES, LEARNING TASKS, AND TESTING IN FOREIGN LANGUAGES

by ERIC W. BAUER

Good testing should be based on effective and well motivated learning. It should be desired by the student rather than hated, and recognised by him as an important learning experience and a means to assess his achievement and progress within the learning system. Tests given at frequent intervals not only measure student progress, but also improve the degree of learning motivation and can be used for analysis and improvement of the teaching program. "Teaching tests" fulfill a function also as "learning tests". In our English Language Programs at Monash University we use diagnostic tests, achievement tests and a great number of weekly "progress tests". The areas tested usually are listening comprehension, structure recognition and production, vocabulary recognition and production, sound-symbol correlation, intonation, stress and juncture, and speaking, reading and writing. In speaking tests, sound production, speed, intonation and word and sentence stress are evaluated in addition to structure and vocabulary. A report on our program will be available shortly. I shall comment on our experiences with these tests in the workshop discussions.

What I would like to do in this paper is not so much to discuss types of tests and the mechanics of testing but to stimulate thought about language learning with a view to new approaches to testing within a learning system.

For this purpose I propose to present a model for learning a second language which I have recently developed on the basis of existing models for taxonomies of learning objectives.

Foreign language learning is something special and as Carroll has pointed out (1962) it is not necessarily dependent on intelligence. It is a complex behavior which is highly successful in individuals who can combine their considerable potential in sensory imagery with abilities in associative recall, in "phonetic memory", in "grammatical sensitivity" and with a high motivation to accept, study and acquire "foreign" or different social codes.

Foreign language learning most certainly must include inductive as well as deductive processes, because cognition and production in the communication act necessitates them and reflective and associative thought involves them. These processes stretch across the areas of all language skills — and are not limited to the one or the other "skill" — because language is an integrative behaviour and our ultimate goal is the "total language experience" although not all individuals can or want to reach this in all instances.

It may be useful to look at language learning as a process which occurs simultaneously on three planes or in three domains: that of "willing" (the *Affective Domain*), that of "thinking" (the *Cognitive Domain*) and that of "doing" (the *Psychomotor Domain*). The first, and in the overall analysis perhaps most important one, deals with attitude, motivation and interest. The second is concerned with all processes of the human mind such as cognition, coding, the formation of *Gestalten*, associations and so forth. The third domain encompasses all activities in language production — whether with the vocal organs, with the hand or through body movements, facial expressions and gestures, etc.

This "tripartite" division is of course only adopted for reasons of convenience in systematization and organization — it never exists in reality in clear-cut divisions but only in a complex web of interconnections and correlations as the human computer creates them in the mysterious "language acquisition device".

The organization of learning objectives in these three domains was developed in a taxonomic model by Bloom (1956) Krathwohl, *et al.* (1968) for learning objectives in subject matter fields other than language. Since foreign language learning is quite different from other types of learning, one might object to an application of such a taxonomy to our field. However it may be interesting to examine this proposition before we reject it. Such an examination should be undertaken with two important provisos: first, that our model is not meant to be prescriptive but only a framework of a descriptive nature as far as it can lend itself to an adequate, comprehensive and economical analysis; and second, that it should be regarded as highly flexible so that adaptations can be made for different types of learners (and groups of learners) and emphases can be ordered according to specific needs.

The model which I am presenting to you was developed for adolescent learners and adults and does not apply to young children and their process of second language learning, although I think it could be adjusted to all age levels.

The interrelation of the three domains, cognitive ability, psychomotor skill and affective behavior can be illustrated by the following graph (p. 184).

The graph illustrates (1) the types of perceptual and motor skill learning on the linguistic surface level, and (2) the types of cognitive abilities learning and affective behaviors on the referential, imagic, and cultural levels.

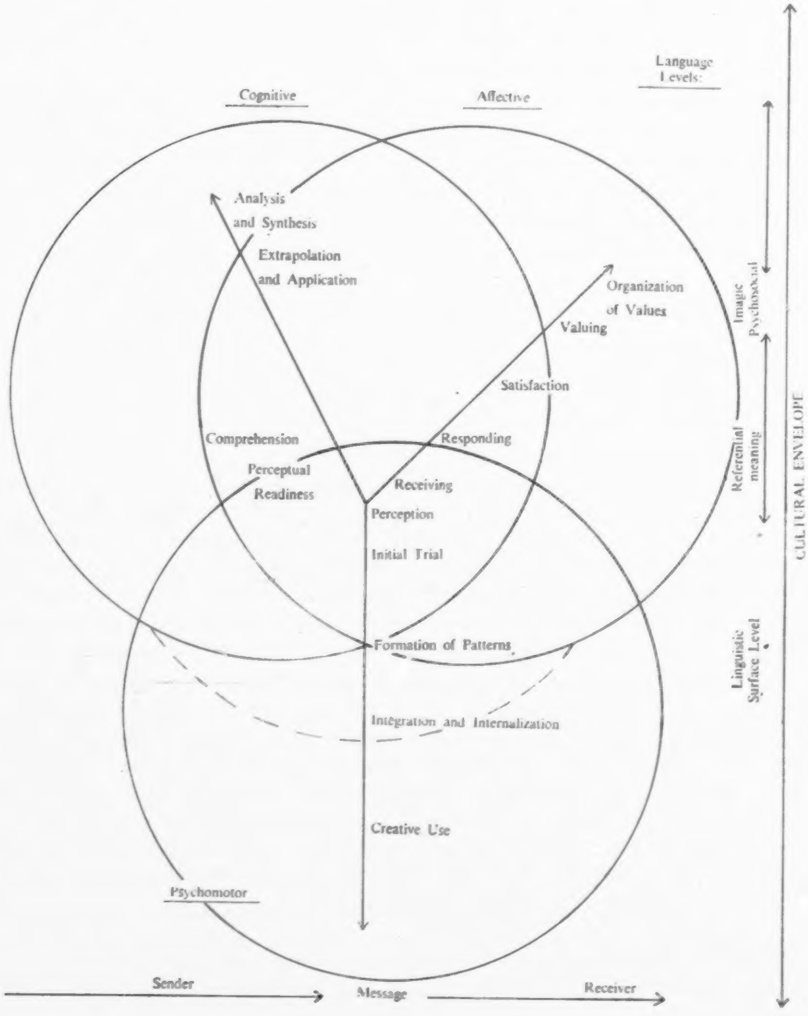
With such a schema it is possible to visualize the various abilities and skills needed for effective language learning which can be identified in terminal criterion behaviors. From these, objectives can be derived and an inventory of testable learning tasks can be established.

In developing the above model I attempted to arrive at a basis for a new unified conception of language learning upon which a more meaningful and more comprehensive hierarchy of objectives can be built, and which may also contribute to greater ease and economy in organizing the learning process. As Bloom (1965, p. 21) says: "Properly used, a taxonomy should provide a very suggestive source of ideas and materials for each worker and should result in many economies in effort." Thus, the taxonomy developed in my study ultimately aims at a better understanding of the learning process through improved means of analysis.

Krathwohl's taxonomy of educational objectives in the affective domain comprises five major categories: receiving, responding, valuing, organization, and characterization by a value or value complex. The cognitive domain according to Bloom comprises six categories: knowledge, comprehension, application, analysis, synthesis, and evaluation. Each of these categories has been carefully analyzed in its application to second language learning. This resulted in some necessary substitutions, extensions, and modifications of categories before specifications could be made for individual objectives in each of these categories.

At first sight these categories appeared to find justification in a hierarchical system of objectives for language learning. However, the categories had to be reinterpreted and reappraised in the light of the source sciences; i.e., in consideration of pertinent studies in research and theory of phycho-linguistics, of socio-linguistic analysis, and of research data in language acquisition. In a unified approach one could define cognitive levels as dealing with major components of competence in language as opposed to psychomotor levels of performance, or better, cognition as opposed to production. Both of these, however, must be integrated into overall learning goals. They must build on each other. Only through constant observance of the inter-dependence within and among domains can taxonomies of objectives and of testable learning tasks become useful and appropriate tools for the analysis of the complex behavior of language learning. The interrelation and co-occurrence of cognitive, affective and psycho-motor behaviors can be regarded as a continuing feature of the learning process. The foreign language student obviously often uses motor responses to manifest affective

THE THREE DOMAINS IN FOREIGN LANGUAGE LEARNING


Cognitive
Affective
Language Levels:
Analysis and Synthesis
Extrapolation and Application
Organization of Values
Valuing
Satisfaction
Comprehension
Responding
Perceptual Readiness
Receiving
Perception
Initial Trial
Formation of Patterns
Integration and Internalization
Creative Use
Psychomotor
Imagic Psychosocial
Referential meaning
Linguistic Surface Level
CULTURAL ENVELOPE
Sender
Message
Receiver

behaviour, cognitive awareness, understanding and application. From overt responses one infers that the student has applied his cognitive abilities in comprehension and application or that he has reached a certain level in the affective domain. We are limited to this assumption because no entirely objective and comprehensive test measures have been developed for all three domains. So-called objective tests inevitably include some degree of the test developer's subjective judgment. In foreign language learning no domain can be assessed while excluding components of the other domains in the medium of language.

An example in cognitive and psycho-motor behaviour may illustrate this interrelation. A motor performance such as using the mother tongue or the foreign language for a response in writing and speaking or for kinesic behaviour should not be regarded as part of a criterion behaviour of cognitive ability. Overt responses to listening comprehension, "discrimination" or "interpretation" may be manifested by checking multiple choice items or answering yes-or-no questions, underlining, etc. However, in either case the psycho-motor response act only partially *manifests* a cognitive "skill" which is *not* a psycho-motor skill but a "skill in the use of the mind" (Henle, 1966, p. 53). Some psycho-motor skill responses and multiple discrimination tests, then, may only be vehicles for teaching and measuring cognitive abilities. They may be affectively rewarded through free access and with unrestricted facilities for repetition until mastered.

The term "skill" should also be differentiated from cognitive ability (Mark, 1969). Skill denotes the acquisition of a habit or level of proficiency in a specific task or group of tasks which ultimately does not necessitate a conscious effort. Contrariwise, cognitive ability involves a conscious thought process and a more general trait of the individual (Fleishman, 1966, p. 147).

This model may be useful for a redefinition of language learning. The term "skill" henceforth could be reserved for the psycho-motor domain only. However, in language learning, combination behaviors represent the ultimate goals which lead the learner to creative expression. Therefore, in the categorization and definition of some objectives, such *combination behaviors* should also be considered in test construction. Precise analysis of the constituents of composite behaviours in verbal learning are still missing since quantitative values or weights have to be assigned each constituent process and combination rules have to be worked out in order to arrive at a synthesis of composite behaviors (Underwood, 1966, p. 498).

Pending the development of such quantitative assignments, we can accept a strategy in deriving objectives in the three domains of language learning which may correct common misconceptions about teaching foreign language "skills" by means of a mechanistic skill approach which ignores or neglects the cognitive and affective components. Skills in foreign language learning really represent basic types of proficiency in the mechanics of language and speech which the sender and the receiver manipulate on the linguistic level (see graph below p. 187). Cognitive abilities in foreign language learning are concerned with understanding, interpreting, extrapolating, analyzing, and synthesizing on the referential, imagic, and cultural envelope level. These levels seem useful for the classification of learning objectives in order to differentiate language skills from abilities.

Since language competency is a complex behavior, the first problem arises with the question for the role of specificity of simple behaviors which lead to complex or more internalized behaviors. The strategy employed in my study basically followed the one suggested by Bloom and his co-workers. It was also described by Banathy in his theory of selection and organization of content (1956). We made every effort to keep the taxonomy neutral with regard to educational principles and philosophies and, as much as possible, objectives from all educational orientations were included in the continuum. The process of selection and classification was that of analysis-synthesis-reduction to practice. In this way we proceeded from a definition of need and purpose as established by the society, the educational system and the learner, to overall educational objectives. Then, in the light of the relevant base sciences such as psycholinguistics, psychology of learning, linguistics, semantics, social psychology and ethno-linguistics, we proceeded to an interpretation of specific-to-language objectives for the adolescent and young adult learner of a foreign language.

It was necessary to proceed in two steps: (1) through an iterative discussion of objectives in foreign language learning, and (2) through an attempt to verbalize the objectives evolving from the discussion-analysis in the form of behavior definitions as suggested by Mager (1962), Tyler (1964), and others.

Each of the groups of objectives was divided into sub-groups of possible, not mandatory, goals which may be needed and which from a practical point of view may occur in the learning continuum. This means that some of the sub-goals may co-occur or be integrated in other goals, although the main groups of objectives are supposed to build on each other in sequence, thus leading from the more specific to the more complex. Since language behavior can be interpreted as a three-level process composed of continua in all three domains of learning, the behaviors specified in the objectives of each domain do in fact largely co-occur, interact, and complement each other. For convenience of analysis, the objectives can be listed in each domain separately. Reference to sequence and interaction with objectives in the other domains can be made in specific instances.

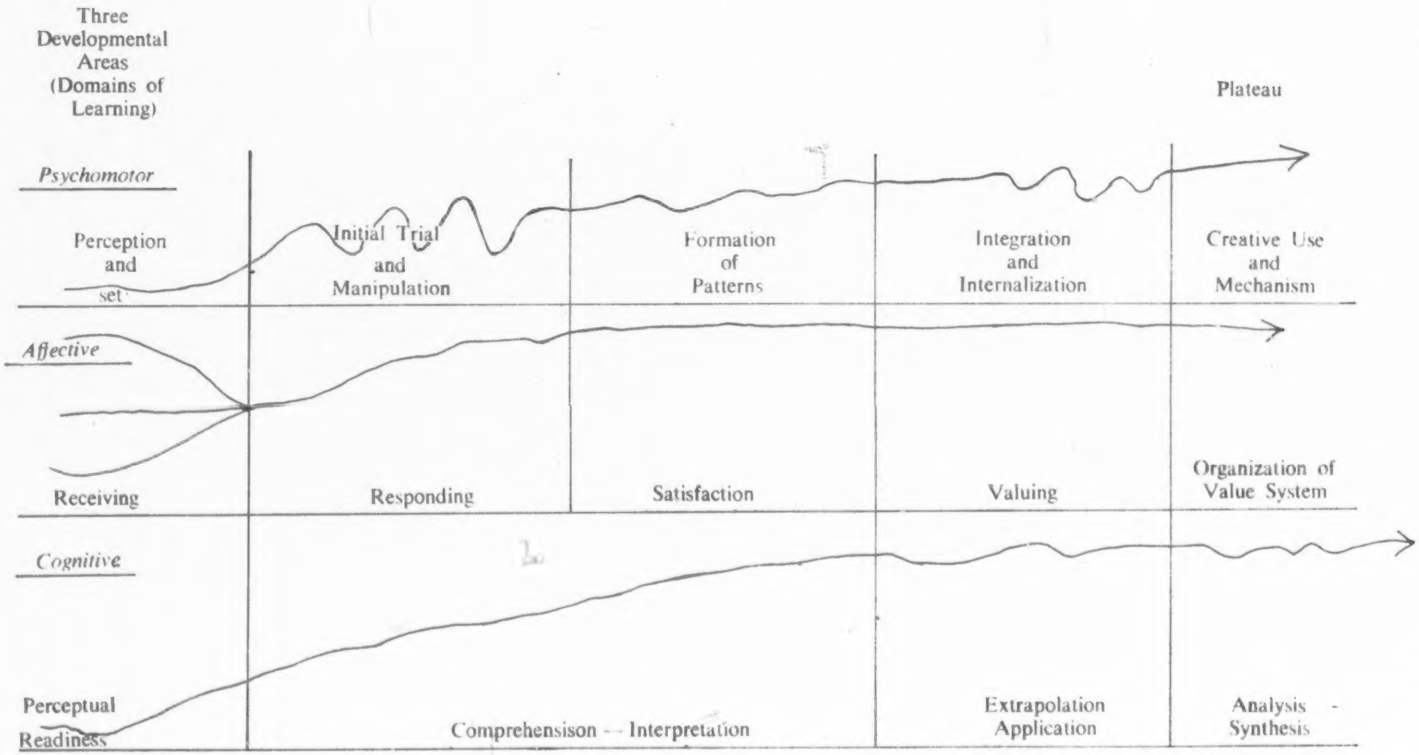
A new classification of objectives is presently being developed and tested at Monash University. We hope it will be of service to the materials developer, and the curriculum-maker in his decision-making process about where to place emphasis and about which sequence to follow. The objectives are adjustable to student variables and variables of program-concentration. They will yield inventories of testable learning tasks.

Specific decisions must be made by the materials developer in selecting objectives for special programs according to special needs and possibilities of the student. In this respect, for example, not all objectives may be valid for all learners in all first year programs. In the graph below you see an example of variables of program concentration as it was worked out for a specific elementary program in foreign language.

The graph shows the interrelation of the developmental areas along continuum curves in the three domains of learning. Their comparative magnitude indicates the emphasis assigned to progressive levels in each domain as the language skills are developed. Each of the five stages of the developmental process of psycho-motor skills represents a category for the classification and specification of objectives.

The model is subject to change and validation. Its purpose is not to present one best solution. It is based on the hypothesis that at every stage of the development of the three components of learning there is possibility for growth and a need for correlation and coordination although emphases may change on various flexible individual scales.

THE DEVELOPMENT OF LANGUAGE LEARNING


Three Developmental Areas (Domains of Learning)
Plateau
Psychomotor
Perception and set
Initial Trial and Manipulation
Formation of Patterns
Integration and Internalization
Creative Use and Mechanism
Affective
Receiving
Responding
Satisfaction
Valuing
Organization of Value System
Cognitive
Perceptual Readiness
Comprehensison — Interpretation
Extrapolation Application
Analysis - Synthesis

Most recently the question has often been raised whether habit-learning is the basis of all second language learning, or whether it should be rejected. While the program-developers in the above specific example could not give a conclusive answer they decided that a one-sided interpretation of skill learning should also be avoid. The question is not whether learning a second language is based on habit-formation or not. It should rather be: when does it take place, for what purpose and to what extent. The important factor, it was pointed out, is the amount of interaction between the psycho-motor and the cognitive and affective behaviours. "Parrot-like" language behavior so often criticized in early audio-lingual programs resulted from an overemphasis assigned to mechanistic motor-activities at the expense of cognitive and affective processes.

It was assumed that the "communication level" could be reached by developing the psycho-motor domain alone through practice in habit-formation and conditioning of S-R processes. This assumption proved to be unrealistic if not impossible. The major handicap really was the lack of balance among the different domains of learning. It is not impossible in a formal school situation to reach the communication level within certain limits, but it is impossible to assume that it can be reached through the development of the psycho-motor domain alone at the expense of cognitive and affective behaviours. From this we may hypothesize that the learner is unable to develop his psycho-motor skills in communication beyond single-item manipulation unless the cognitive and affective levels are developed to an appropriate degree. The degree of attainment in the psycho-motor domain may increase only after a latency period with initially much higher attainment levels in the other two domains.

Since language learning aims at long-range retention, we should be seriously concerned with recall and retention of cognitive-psycho-motor behavior. In each of the four skills a solid foundation for recall and retention can be laid in the comprehension-interpretation phase of the learning process during which the formation, organization and some manipulation of motor skills takes place. Ideally, in this phase, the learner also develops an active interest and a warm, positive attitude towards the subject, the culture, and the learning system. He concentrates his efforts on firm formation of learning sets in all three domains and makes frequent attempts to interpret new material, while at the same time being open to refine or change his "guesses" or approximations which may prove to be inaccurate. The learner's open attitude and willingness to respond and to try may well be decisive for integration, patterning, and retention of motor behaviors and concepts. Once efforts at interpretation have begun, the learner can be reinforced and rewarded when he encounters the material again in context and when his interpretation is confirmed as being correct (for example through "guided learning experiences" in self-instructional programs). This material, then, becomes more readily available not only for listening and reading, but also for speaking and writing, since it has been derived in a conscious thought process associated with psycho-motor clues within appropriate context. Neither the number of occurrences nor the amount of practice is decisive. Rather, it is the strength of impression and the contiguity as the learner perceives it, the development of his "feel" through a willing mind, that is decisive.

Before approaching the classification of goals in the psycho-motor domain, we may want to examine various performance criteria as observable and measureable for individual goals.

Performance criteria for the psycho-motor skills

Performance criteria for each of the categories in the psycho-motor domain can be established on the basis of the following factors or combinations of these:

(1) *reaction time* for making a response.
(2) *speed* of delivery of a response.
(3) *accuracy* of a response.
(4) *meaningfulness* of a response or clarity of intent.
(5) *selection of the appropriate cue* for the orientation of a response in complex tasks.
(6) *steadiness* of one or several of the above factors.
(7) *meaningful coordination* of a series of responses.

It may be noted that factors 1) to 3) above constitute the "automaticity" of a response; factors 1), 3), 4) and 5) are related to the comprehension of the task and the material, and factors 5) to 7) indicate integration of a more complex psycho-motor task. If all factors are combined and developed to a maximum, optimum integration has been reached in performance.

Evidence for the functional value of these factors in verbal learning can be found especially in psychological research. (Fleishman, 1966; Underwood, 1966) Observations about motor skill learning and verbal learning refer to reaction time, perception and search processes, coordination of individually learned skills, stimulus selection, and continuous performance. These may also be of interest to second language learning and testing.

In research on language teaching, workers such as Marty (1966) stress accuracy in reading aloud, spelling, written expression, oral expression, and comprehension, as well as oral fluency and reaction speed as particular goals. In language testing, specifications can be made and levels of performance can be established especially for *speed* (rate) of delivery in speaking, listening, and reading: *accuracy* in all four skills: *appropriateness* and *meaningfulness* of a response in listening and reading comprehension and speaking, and *meaningfulness* in speaking or writing utterances beyond the sentence level (paragraphs). However, so far, not much consideration has been given to a recognition of all the above factors (1) in all four language skills, and (2) for verbal as well as non-verbal responses (action-responses or kinesthetic responses).

Factor 5), *selection of the appropriate cue,* has been too narrowly conceived in our teaching and testing procedures. Cue selection for orientation of a response, which the learner makes through sensory perception and cognition, can be regarded as an objective and criterion in itself and need not be contaminated through the additional difficulty of a motor response goal. For instance, in audio-comprehension tasks, the learner should get recognition for appropriate responses in his native tongue. He should also be given credit for partial fulfilment of a communication task if he is able to recognize signals of various audio-visual "learning modes" and if he can prove that he could react appropriately in his native language. (viz: Testing devices as provided in the developmental stages of the Carroll-Sapon MLAT).

In my research I have experienced the value of recognizing this factor in all three domains of learning: *attitudes* (student's confidence and motivation), *cognitive understanding* (the student's train of thought was not interrupted), and *psycho-motor* (although the student sometimes gave a response in English, he afterwards also tried to respond in the foreign language).

It is important to note that the identification of the above factors has a great potential in considering individual differences by allowing various levels of performance in varying combinations because perception in second language learning relates to the student's sensitivity to different visual, auditory, or kinesthetic cues. The pre-requisite to any verbal psycho-motor activity is the cognitive aspect of perceiving the task, or the identification of sensory stimuli; e.g., different grammatical signals not used in the native language, phonemic and melodic contrasts. In the case of a stimulus composed of more than one element, the student has to perceive and identify the stimulus elements which are relevant for the task at hand. He also has to make the necessary neuro-muscular or attitudinal adjustment necessary for carrying out the verbal act.

References

Banathy, B. H. *A theory of selection and organization of content in foreign language curricula.* Unpublished doctoral dissertation, Univ. of California at Berkeley, 1966.

Banathy, B. H. "The systems approach". *MLJ*, May, 1967, pp. 281–289.

Bloom, B. S. (ed.) *Taxonomy of Education Objectives. The Classification of Educational Goals. Handbook I: Cognitive Domain.* New York: David McKay, 1956.

Carroll, J. B. "The prediction of success in intensive foreign language training". In R. Glaser (ed.), *Training Research and Education.* Pittsburgh: Univ. of Pittsburgh Press, 1962, pp. 87–136.

Carroll, J. B. "Psychology, research in foreign language teaching: the last five years". In Robert J. Mead, Jr. (ed.), *Language Teaching: Broader Contexts* Northeast Conference Reports, 1966, pp. 12–42.

Fleishman, E. A. "The ability-skill distinction". In E. A. Biloudeau (ed.), *Acquisition of Skills.* New York: Academic Press, 1966, pp. 152–157.

Henle, M. "Cognitive Skills". In J. Bruner (ed.), *Learning about Learning — A Conference Report* — Washington, D. C.: U.S. Dept. of Health, Educ. and Welfare, 1966.

Krathwohl, D. R., B. S. Bloom, B. B. Masia. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: Affective Domain.* New York: David McKay Co., 1964.

Mager, R. F. *Preparing Instructional Objectives.* Palo Alfo, Calif.: Fearon Publ., 1962.

Marty, F. L. *Language Laboratory Learning.* Wellesley, Mass.: Audio-Visual Publications, 1960.

Mark, G. M. A. *Improving Teaching Skills Classroom Management for Effective Instruction.* Unit Five of English Unit Series Chicago: Science Res. Associates, 1969.

Tyler, R. W. "Some persistent questions on the defining of objectives". In C. M. Luidvall (ed.) *Defining Educational Objectives.* Pittsburgh: Univ. of Pittsburgh Press, 1964.

Underwood, B. J. "Motor skills learning and verbal learning: some observations". In E. A. Biloudeau (ed.), *Acquisition of Skills.* New York: Academic Press, 1966.

WORKSHOP REPORTS

WORKSHOP 1 — TESTING LISTENING SKILLS

RALPH P. BARRETT

Workshop 1 was assigned the problem of testing listening skills. The first meeting on Wednesday afternoon was devoted to a discussion of the nature of these skills and the problems involved in evaluating them. A short time was devoted to an explanation of the different functions of the proficiency test, the diagnostic test, and the achievement test, noting the range which each is designed to evaluate on the full scale of language proficiency from zero knowledge to native speaker competence. We also talked about problems of validity and the use of hybrid versus pure tests, as well as problems of reliability with respect to the use of certain objective (recognition) item formats versus subjective (production) item formats. At the end of the first session, a survey was conducted to discover what the workshop participants wished to achieve in the four remaining sessions.

Out of this survey and our initial discussions came several decisions and guidelines; it was decided that (1) we would attempt to produce an *inventory of techniques* for testing listening skills rather than a full test itself, (2) we could limit these techniques as much as possible to objective-type multiple-choice or partial production responses, in order to eliminate subjectivity in grading, and (3) we agreed to produce testing techniques that would be appropriate for students at three general levels of English proficiency: elementary, intermediate, and advanced.

We also agreed *on what aspects of the listening skills* we would attempt to test. We reasoned that *if* an English student does not comprehend what he hears spoken to him by a native English speaker, it might be for three distinct reasons. First, he might not be able to discriminate the phonemes of English well enough to recognize words that he would know in their written form. Second, he might be able to "hear" the words but still be unfamiliar with their meaning, especially when used colloquially. Third, he might know what the content words mean and still miss the total meaning because he is unfamiliar with the underlying meaning of the grammatical structure. It was decided then to test four areas of listening comprehension: (1) "global" comprehension (that is, the total comprehension of a sentence or passage or conversation), (2) comprehension of vocabulary in a structurally simple spoken context, (3) comprehension of structural meaning in spoken contexts with minimal vocabulary difficulty, and (4) the discrimination of English phonemes, including perception of stress and intonation patterns.

The workshop participants chose which of the four areas they wished to work in, and the remaining two days were devoted completely to the production of testing techniques. So much material was forthcoming that it is impossible to describe it adequately in this brief report. Each group succeeded admirably in developing a set of useful and creative item formats for testing different proficiency levels. The use of pictures as partial cues or as answer choices by all groups was particularly interesting, suggesting a renaissance in the use of this long-neglected testing technique.

Two general observations about the material might be of interest. First, the participants were especially interested in the testing of the ability to perceive and to assign meaning to stressed and unstressed syllables and to intonation patterns, and a good number of techniques were designed for this purpose. Second, few techniques were developed for testing vocabulary at the advanced level, since it was decided that high-level vocabulary was best evaluated by using the "global" method-comprehension testing of combined structural and vocabulary problems.

Several recommendations regarding listening comprehension testing came out of the workshop situation. The most important was probably the suggestion that before a listening test is administered, the items should be checked by native English speakers representing both British and American dialects. This is to guard against items of vocabulary or grammar which might be appropriate for speakers of one dialect but not of the other. Another safeguard is the use of the tape-recorder to administer the listening test, thereby insuring against those inadvertent slips of the tongue and annoying hesitations which can disrupt an otherwise smooth presentation of the oral cues. Of course, in the standardized testing situation where two or more groups are given the same listening test, it is highly recommended that the spoken cues be recorded to insure a uniform presentation to all of the groups.

Before concluding, I would like to give special thanks to Mr. David Cobb and Mr. Jonathan Trench, both lecturers in English at the Asian Institute of Technology in Bangkok, for generously sharing with us their recent work in testing listening comprehension and for assisting so ably in their respective work groups. And of course, my thanks to all of the workshop participants for their splendid cooperation and their hard work which made it possible for us to achieve our workshop goals.

WORKSHOP 2 — TESTING ORAL SKILLS

CHARLES R. SHEEHAN

During the first workshop session the members were asked to give a brief description of the types of oral testing they had done or were required to do. As a result of this session, it was decided to work on oral proficiency tests. At the second meeting the oral proficiency test of the Foreign Service Institute, U.S. Department of State, was explained and group members prepared an oral test designed to test the behavioral goals of the S–1 level or some aspect thereof. The members wrote these tests in the respective native languages represented by the group. Three of these tests were administered by members of the group to other members of the group. The three tests administered were Chinese, Thai, and English. Following the tests was a discussion on the format of the tests, what it was designed to test, and an attempt was made to assign an FSI S-rating. During this discussion period. it became evident that while the majority was interested in overall proficiency testing, it was not the type of testing they were required to do in their roles as English teachers. The rest of the session was spent establishing sub-groups for work on more specific topics in terms of members' interests and needs. This resulted in the establishment of four sub-groups: (1) Techniques for Oral Diagnostic Testing in the Classroom; (2) How to Define the Terminal Behavior Desired from a Student in the Use of English, then decide if the course is geared to teach that behavior and does the testing test that behavior; (3) Simple Devices to meet Classroom Need for Testing Oral Production of English and the discussion of a specific oral proficiency test — its administration and how it can be used as a teaching device. The four sub-groups were lead by Messrs. Sheehan, FSI, American Embassy, Bangkok; Bloomfield, Chong Chi College, Hong Kong; Botsman, New Guinea Institute of Technology; Klein, Defense Language Institute, U.S. Department of Defense, Naval Advisory Group, Bangkok. Mr. Klein's group was made up of two members: Dr. Leocadio, Philippines; Dr. Lim, Singapore. During their session they were able to establish proficiency level requirements for specific professions, technical fields and non-technical fields in their respective countries. This was done by applying Mr Klein's method of evaluating his oral interview test. Specific percentages were assigned to various job-level English requirements.

Due to the nature of topics in other sub-groups and individual needs, only observations can be derived from their sessions.

The following observations were made by the workshop group members:

1. That some English teachers did not realise that the continuing assessment of student oral production through the use of guided conversation, transformation and other pattern practice, and oral response to pictures could be classified as oral testing. There seemed to be confusion between formal and informal testing.
2. There seemed to be an intuitive need on the part of language teachers to give some type of oral examination as normal classroom procedure even if the school itself has no requirement for formalized oral testing *per se*.

3. Some countries have an oral interview or overall proficiency test for entrance into Teacher Training College, while others do not.
4. It seemed apparent that some school systems teach English as a tool subject at lower developmental levels with emphasis on reading skills. This has caused difficulty at higher levels such as teacher training colleges and some technical schools in which an active knowledge of English is requisite for success.
5. Non-systematic oral testing seems to contribute to lack of student motivation to study oral English in some countries. No recommendations were made.

WORKSHOP 3 — TESTING READING SKILLS

B. DEANE SMITH

Our workshop devoted its initial sessions to general discussion of the topic, with the purpose of assessing the interests of the group and deciding on specific objectives of our activity. We found that there were two categories of interest: one group of participants was especially interested in the initial stages of reading; another large group was interested in more advanced, developmental, stages of reading. All the participants wanted to do something practical, and we agreed to devote our efforts to the evaluation and construction of reading test items.

For evaluation purposes, the group developed the following list of criteria:

(*a*) What skill does the item test?
(*b*) What kind of response is required by the item?
(*c*) Does the item really test the desired skill?
(*d*) Is the item difficult or easy to construct?
(*e*) Is the item difficult or easy to score?
(*f*) Is the item suitable for one level or another?

As a working procedure the group divided into four sub-groups, one devoting its attention to the initial stages of reading, three to the developmental stages. There was general agreement that the purpose was not to construct a model reading test, but rather to examine, construct and evaluate a wide range of types of test items.

The sub-group working on the initial stages of reading established a frame of reference for evaluation, based on eight skills as follows:

(*a*) Sound-to-letter transfer.
(*b*) Sight recognition of function items.
(*c*) Passive recognition of grammatical structure.
(*d*) Productive use of the above.
(*e*) Word meaning.
(*f*) Recognition of significant details in context.
(*g*) Recognition of main ideas of a passage.
(*h*) Correct inferences or interpretation.

This group produced an annotated list of test items, which is, however, far too lengthy and detailed for inclusion here. I would like to express my appreciation to Peace Corps Volunteer Gary Arrowsmith for his initiative in the work of this sub-group.

The other sub-groups chose a common reading passage for reference. One group developed a simplified version of the passage for its use, the others worked with the unsimplified version.

The results of our two and a half days of work were not a series of recommendations to be offered to the plenary session. The workshop did not consider this as their objective. A large amount of production of test items and evaluative annotation of items was produced. This cannot be presented here, but can be made available.

The real value of our association of effort lay in coming to grips with the problems of preparing and evaluating test items, and the function of these items in the evaluation of reading.

WORKSHOP 4 — TESTING WRITING SKILLS

OLIVER SEET BENG HEAN

The theme of this workshop was to explore the possible question types to test students' writing skill. Test items were to include vocabulary, grammar and the mechanics of writing (at the objective level) leading to productive techniques in composition and guided composition. Three stages were identified for testing, representing three ability levels.

Various types of discrete item tests were reviewed; these included recognition, production and partial production techniques. The functions, limitations and construction of different forms of multiple choice, completion and matching questions were fully discussed as well as techniques of increasing the degree of difficulty for various levels of attainment. The group also considered different forms of contextualized objective tests. It was noted, however, that these types of questions did not demand a productive use of the language and therefore were suitable only for beginning students. At the upper levels of attainment these had to be complemented by types of tests requiring productive skills. The participants then divided into three groups for the development of test items in objective form for use in the region. The following areas were considered — Grammar, Vocabulary and Mechanics. The materials produced by each group will hopefully be made available to the participants in this seminar. They are too substantial to be scrutinized in this brief report.

In discussion the complete group noted the difficulty of free composition as a test item but still felt that it was inappropriate to reject it entirely as a test of writing skill.

The objectives of composition writing were formulated as follows:

(1) to test active usage of mechanics and grammatical, semantic and lexical items learnt;

(2) to demonstrate coherence;

(3) to elicit continuous prose;

(4) to test knowledge of phrases and forms peculiar to the written language.

It was pointed out that discrete item testing is relatively passive and therefore continuous writing was necessary at some point in any teaching programme. Two possible approaches to scoring were discussed. In the analytical approach, it is necessary to set up a check list and the group agreed on a check list of the following items:

(1) Mechanics (which would include paragraphing, capitalization, punctuation and spelling);

(2) Grammar (which would include concord of number, tense and pronouns, preposition, articles, adverbs and word order);

(3) Semantics (which would include variety and topic appropriateness);

(4) Lexicon (word count);

(5) Content and Form (which would include points of information, organisation and sequencing).

This is clearly a complex form of marking and only feasible where a small number of scripts have to be marked.

The second possibility would be scoring by impression. In this form of scoring, the importance of achieving a consistent standard of marking was stressed and the various methods of moderation used in different parts of the world were reviewed. Different forms of multiple scoring of themes for greater reliability and the problem of preventing a shift in consistency in individual scoring over a length of time were also reviewed.

It was agreed that the analytical approach to scoring has not proved to be more reliable than impression scoring and that only partial objectivity at best could be attained in either form of scoring. A greater degree of objectivity could be attained, however, in multiple scoring system and where examiners took pains to maintain consistency by selecting models from the scripts marked and occasionally returning to these models.

Discussion then centred on the relative importance of (*a*) content and (*b*) use of language, but it was difficult to reach any agreement on this in the presentation of free composition.

Finally in the matter of techniques, several suggestions were made for eliciting continuous writing by guided/controlled prompts and it was felt that an analytic assessment or evaluation could be made here which was perhaps not feasible in free composition where large numbers of scripts were involved.

WORKSHOP 5 — TESTING SELECTED AREAS OF ENGLISH GRAMMAR

ERIC W. BAUER

Delegates from six SEAMEO countries took part in the workshop; they represented Indonesia, Laos, Malaysia, Philippines, Thailand and Vietnam. Before the first meeting of the group the leaders of Groups 5 and 6 agreed to arrange for an initial joint session in which both chairmen could present their working programmes. In this joint session the leader of Group 5 presented a paper "Objectives, Learning Tasks and Testing" and the Chairman of Group 6 outlined the programme of his workshop.

The purpose of the workshop was (1) to discuss the role of grammar in the Teaching and Testing of English, (with special reference to the aural-oral skills) (2) to specify ways and techniques to test selected problems in grammar which pose major difficulties in English Language Programmes.

It was agreed that testing grammatical areas only serves as a tool in language development towards effective recognition *and* expression. Grammar should be presented and tested in a realistic and purposeful way so that the student's motivation can be sustained. Mastery of grammatical features is essential for listening and reading comprehension as well as for speaking and writing. Only if grammar is tested across language skills will we have a good idea, in how far the student has internalised it, of what a student can do. Unfortunately we do not have good grammar tests in all these areas.

It was pointed out that different levels and target groups have different objectives in grammar. The members of the group defined variations of specific grammatical goals as they are applicable to their target groups of learners. Although it was realized that different objectives for cognition and production may be established in some areas of grammar, minimal standards have to be maintained. Therefore, the recent suggestions in the literature of "semi-grammaticalness", must be examined with great caution and cannot presently be accepted. "Communicative productivity" is an important goal and language is highly redundant in many features. Yet, the establishment of good standards and the specification of objectives in phonology and grammar, it was agreed, are the foremost objectives in teaching and learning English in South East Asia.

The Workshop group then delineated various goal areas for intermediate and advanced language learning which should be tested, such as

(1) grammar review and extension,
(2) vocabulary review and extension,
(3) recognition of cultural patterns,
(4) specific reading skills, especially making inferences, "gisting",
(5) specific listening skills (also in connection with the reading skills), especially those of "associated recall" and "interpretation" (in terms of the special meaning of "interpretation of context"), extrapolation, gisting, "chunking", and note taking,
(6) the speaking skills.

Considerable time and effort was devoted to discussing the testing of grammar in listening comprehension and oral expression. The chairman reported on his research and testing programmes in the U.S.A. and in Australia and in particular about methods of oral testing. Three types of oral tests were discussed and examples of test items were presented. Special reference was made to the interview method. Responses in a foreign language in guided or free communication tests should be as economical as possible, i.e. to the point, with a minimum of verbiage and the correct grammatical and idiomatic choices to be made in a minimum of time. These factors ought to be considered in the scoring system. It was emphasized that various skills can be tested and scored separately in oral examinations.

It was felt necessary to discuss techniques which could be used in testing before specific areas of grammar were selected by the group. This was done to establish terms of reference with which all members of the group were familiar. The chairman outlined the major testing techniques for grammar and Dr. Murphy reported on his work in Indonesia. In the group discussion, specific examples of teaching and testing items were worked out, such as applicable in transform and transformation, substitution and completion, correlation, combination, expansion, response and sequence techniques.

The group then defined the areas of grammar that they considered most important in their teaching. These were: tenses, word-building, agreement, articles, pronouns, two and three part verbs and prepositions.

Each of these areas was discussed in relation to testing techniques which are most appropriate for them.

There was considerable discussion on the problem of tense, since this poses the greatest difficulty in the countries of this area. Several of the languages e.g. Thai, Malay, have no grammatical tense markers as in English but use reduplication or lexical time signals. The members of the group agreed that many students can learn the forms of the tenses but have difficulty in fully understanding their use. Testing the understanding of forms through multiple choice items or other devices is not sufficient. One also has to evaluate production, i.e. effective application of grammatical knowledge in communication, or application in written expression. Several techniques of testing *understanding* (i.e. the cognitive processes of comprehension, interpretation, extrapolation) and *sensitivity to grammatical meanings* in the contrastive situation were discussed, and a model was presented by the chairman. It was agreed that the best way to teach and to test the tenses was in close association with the situational context. Copies of two articles were distributed and practical applications of some suggestions contained therein were discussed.[1]

Some discussions dealt with the problem of teaching and testing the correct use of articles and of agreement. Dr. Murphy outlined special techniques in testing "agreement" and the participants reported on their work with specific target groups and prepared suggestions for grammatical goals.

Questions about various approaches to present and teach grammatical problems and to evaluate their mastery were discussed. Two members of the group raised the question of the extent to which recent generative transformational analyses of modern English could be used in teaching (and testing). The chairman pointed out that the sequence and the approach in the learning process (and in the establishment of testable learning tasks) were first of all determined by ease of learning and economy and secondly by grammatical criteria based on the logic of descriptive and explanatory analysis presented by linguistic theory. "Deep structure analysis" as far as it can be applied in teaching (and testing) can be demonstrated and generated by the student in meaningful and motivating context.

---

[1] Bauer, E. W. "Approaches to a redefinition of Language." *Contact*, Publ. de la Feder des Prof. de Languages Vivantes, Dec. 1966, 9, pp 11–18.

Chomsky's model may also be acceptable as a source for foreign language teaching because (1) it explains native language learning, (2) it establishes one general base of description, (3) it preserves the creative-symbolic-artistic nature of language, and (4) it establishes a basis for the social task of language learning, namely, to expand one's own knowledge about one's mother tongue in learning a foreign language just as one expands one's own culture by learning about foreign cultures.

Since the scientific description of the native language in modern linguistics is based on an *inference model* of rule-governed behavior, one can establish a *criterion behavior*, which may be broken into graded sub-goals for limited rule-governed behaviors. However, linguistic description of a native language system represents only an analysis of data. For second language learning, these data must be compared and contrasted with the data for the native language system and converted, broken down, limited, and organized into a pedagogical grammar for the learner according to *the purposes and objectives* of the language programme.

During the final sessions the group concentrated on a discussion of the role of grammar. The chairman stressed the need to develop a balanced approach to foreign language teaching, so that comprehension or cognitive goals as well as production goals can be established, and thinking and understanding can be furthered. It was demonstrated how objectives can be established and specified in the three domains of language learning: the affective, cognitive and psycho-motor domains. Once these objectives have been properly specified we can build our teaching and testing programme on these. It was pointed out that a list of such specifications can also be a valuable tool in evaluating whether textbooks and instructional programmes and testing programmes meet the standards laid down in the objectives. Certain criteria for the development of teaching and learning systems can be established to this end.

In the last session a number of recommendations were made by the group to be presented to the Seminar. They dealt with more general problems. It was recommended that:

1. Clear objectives be specified for the various levels from elementary school through to university.
2. These specifications be made for the various domains of language learning.
3. According to these specifications, placement tests be constructed.
4. At the secondary level, four achievement tests be given each semester according to specified objectives.
5. Special types of oral tests be constructed. There should be teachers' manuals that give special examples of speaking tests and how to give them.
6. The curriculum provide for frequent re-entry of the essential learning tasks in grammar. ("Re-cycling" and recombination should be carefully considered).
7. More in-service courses for teachers be provided and of a longer duration. They should be conducted over a minimum of six to eight weeks and they should be intensive (e.g. five hours a day). Present courses are considered insufficient.

8. These courses deal with methodology (teaching and testing techniques) and they also provide for intensive practice in the speaking skills.
9. Arts Graduates who are going to teach English be required to pass a qualifying examination for a teacher's certificate.
10. A professional association of teachers of English be set up which could help safeguard professional standards and make contact with similar associations in other countries, e.g. the National Association of Teachers of English, the Modern Language Association (U.K. and U.S.A.) and TESOL (U.S.A.)
11. Supplementary teaching materials for English be produced, (including audio-visual materials).
12. There be more co-operation ("articulation") between primary, secondary and tertiary levels of instruction in English in the development of instructional materials.

I would like to thank all members of the group, and particulary the delegates from Thailand, for their enthusiasm and good working spirit which enabled us to conduct a very fruitful and enjoyable workshop.

Special thanks also go to Dr. Murphy, Dr. Tirtopramono and Dr. Leocadio for their most valuable contributions.

WORKSHOP 6 — SETTING REALISTIC BEHAVIORAL OBJECTIVES FOR DETERMINING PROGRESS IN LANGUAGE MASTERY AT VARIOUS LEVELS

RICHARD NOSS

SEAMEO countries represented in our workshop were Indonesia, Malaysia, the Philippines, Singapore, and Thailand. This is a summary of group discussions, and also of individual reports and recommendations made by teams of representatives of the five countries, with the help of volunteer group members from outside the region. Since the longer reports are full of valuable information and some stimulating new suggestions, and time does not permit our presenting them here, we hope that RELC will eventually make them accessible to you.

The Workshop held its first session jointly with Group 5, and heard Prof. Bauer's paper before setting its agenda. The next three sessions were devoted to discussion, and the final one to the writing of reports.

1. The first task was for the five countries to inform us of their educational and other objectives in English teaching, their actual course syllabi, and their tests, both internal and external. The focus of this exercise was on the relationships among the three factors: objectives, syllabi, and tests.

As to objectives, although Singapore presses for actual bilingualism, and the Philippines sees among its final aims a much more comprehensive use of English in such things as internal and external communication, government administration, and most of its general and technical education, all five countries agreed on two major objectives, whether explicitly stated or not:

(1) As the vehicle for further study at university level and for international communication.

(2) As a means to social advancement, prestige, and employment opportunities.

As for syllabi and tests, again with the qualification of Singapore, where English is the most frequent medium of instruction, and the Philippines, where it becomes so after third grade, this concern for English as a means to both academic and commercial advancement is reflected in the number of hours given to its study, and in the frequency with which the progress is examined through intermediate testing.

By the final year of school, students are found taking at least six hours of English per week, regardless of their choice in other subjects. Moreover, straight language study is often further extended to the study of literature, and in the Philippines, for instance, to preparation for journalism and other mass media professions.

Unit and periodical tests occur as frequently as six times in a school year, with term examinations occurring as often as twice a year at all levels. Both Malaysia and Singapore employ external examinations such as Cambridge School and Higher School Certificates, while in the other countries there is widespread use of TOESL, TOEFL, and other overseas examinations in connection with study abroad, and Civil Service tests for eligibility in government service.

2. In order to classify behavioral objectives, the group then turned to the study of samples of previous attempts to define standards of language proficiency, such as the U.S. Foreign Service definitions of speaking and reading proficiency on a scale of zero to five. It became clear that such definitions, to avoid circularity, must not specify products of linguistic analysis such as grammar, pronunciation, or vocabulary, but should concentrate on specifying total linguistic behavior in one or more of the four basic skills.

Further, it was agreed that such standards should be considered first of all *as* objectives, whether teachable or not, whether testable or not. In other words, if it is considered desirable, in a given school system at a given level, for students to be able to take notes on a lecture delivered at normal speed by a native speaker, this should be specified as an objective at that level. The job of the course designer or the test constructor would then be to tailor his work to an aggregate of such objectives, to whatever degree possible.

After a preliminary attempt to list all possible behavioral objectives, and rank them in order of difficulty, the group soon discovered that such a list could be almost infinitely fragmented and refined. It was decided, therefore, to construct a classification device instead, and see if all real-life objectives would fit into it somewhere. The device turned out to be a sort of grid, with the four traditional skills as one dimension, with variables to be plotted into it in one or more of six other dimensions. These dimensions are as follows:

(1) Difficulty of content.

(2) Distortion (noise in spoken language, legibility in written).

(3) Speed.

(4) Dialects (including native and non-native speech) or variant forms (including older forms of the written language).

(5) Register and Style.

(6) Degree of comprehension (receptive skills) or appropriateness of expression (productive skills).

It was found that most actual behavioral objectives could be accommodated by this grid, except for a few involving language-related skills such as summarizing, outlining, note-taking, paraphrasing, and skimming. Interestingly, however, translating and interpreting were rejected as involving more than performance level in a single language.

Most important, it was realized by the group that even such apparently unambiguous terms as "total mastery" must almost always be further refined in behavioral terms. Granting that a university entrant, for example, has control of English comparable to his native-speaker counterpart, say in English, we must still ask these questions: At what level of content difficulty? With how much allowable distortion? At what speed? In what forms and styles of English? With what degree of comprehension? These are, indeed, the same questions we would ask about the native speaker of English who is to enter the university.

3. The Singapore, Malaysia and Philippines delegations then went on to prepare individual statements of desirable objectives in their respective educational systems at various critical levels. In the case of the first two countries, we now saw this had to be done twice: once for the English-medium schools and once for the non-English-medium Schools. Thailand and Indonesia yielded the privilege of submitting their recommendations on this subject so that a summary would be quicker to achieve.

The collective recommendations of the three TESL countries, although far more specific than any officially stated objectives, contain some surprises independent of that difference. The emphasis in the English- medium schools is still primarily on preparing students in those specific skills which they will immediately need to get the most out of their other subjects, and eventually, out of higher education. But in addition, there was a real concern for preparing those students who will fail or drop out of school by teaching them skills oriented toward employment, rather than toward further education.

In the non-English medium schools, the recommended objectives tended to be far less ambitious than the official objectives, perhaps reflecting a realization that the goal of true bilingualism is unattainable without the reinforcement of English in other subjects, or at home, even when the student is able to finish school. Again, there was a considerable emphasis on job-related skills.

4. As to tests, the recommendations were nearly unanimous, at least in character, and in the context of our discussions and the revised objectives. Among the most important are probably the following:

(1) Higher standards in oral testing, especially at critical progression stages such as school leaving, and especially for English teacher candidates.

(2) Greater frequency of oral testing, with recording equipment wherever possible, if only as a means of convincing students that the oral part of the course is important.

(3) Objective written tests on total language behavior, rather than discrete item testing, especially at lower levels.

(4) Tests of continuous prose writing (essay-type) to be required at higher levels, or wherever administratively feasible. (There was doubt, however, that such tests could really be standardized.)

(5) University entrance examinations which are objective but much more difficult than they currently are.

Finally, it was the general feeling of the group that educational objectives must be made much more specific, and preferably stated in behavioral terms. These refined objectives, in turn, should determine *both* the syllabi *and* the internal tests, and in fact the normal interplay between syllabus and test should be encouraged, rather than allowing one to determine the other. External examinations, on the other hand, should not be deliberately prepared for, except when sanctioned by the Ministry.

WORKSHOP 7 — TESTING FOREIGN LANGUAGE APTITUDE

JOHN B. CARROLL

(Report presented by Paul I. Aiken, Jr.)

The Foreign Language Aptitude Testing Workshop was able to bring together participants from five of seven SEAMEO countries — Laos, Malaysia, Singapore, Thailand and Vietnam; and among them were a number of our co-workers already actively engaged in a consideration of the question of foreign language aptitude testing in their countries. All in all, it was a very stimulating workshop, covering the central questions concerning the nature of a foreign language aptitude test; of developing and using aptitude tests in each of the countries represented; how such tests could be developed if it were found desirable to do so; what problems might be encountered in trying to use the tests; and not of least importance, how and under whose auspices such tests would be constructed. Since our individual workshop sessions overlapped a great deal in considering these five areas, it will be clearer, I think, to treat each of the problems separately rather than to give a detailed account of each of the individual sessions themselves.

For those of you who heard or read Dr. Carroll's plenary session address, the nature of foreign language aptitude tests – particularly of the Modern Language Aptitude Test (MLAT) — is generally understood. To review the definition briefly, a foreign language aptitude test, unlike a proficiency or an achievement test, is designed to measure *competence potential* and to predict future learning performance on the basis of what an individual can do at the moment. No important assumptions need be made about prior language training (or lack of it) in arriving at a judgement of such potential. In developing the original MLAT battery, the first step was a task analysis of what types of performance are to be predicted. This ran to as many as 25–30 sub-tests. Each hypothesized ability had to be tried out with a concurrent coevaluation method followed by a statistical sorting out of both those subtests which didn't accomplish their goals and an ordinary item analysis and validation with those remaining.

The workshop used, as a basis for discussion, the English language version of the MLAT; and our members actually took the test so that we could become more familiar with that test's contents (and, incidentally, discover our own strengths and weaknesses!). You might be interested in our scores — which ranged from the 95th down to the 3rd percentile — though since this version of the test was designed for native speakers of English — I think we might take those low percentile ranks with a few grains of salt. After taking and scoring the test for ourselves, the workshop session participants went on to discuss the various components of the test — namely (1) number learning (2) phonetic script (3) spelling clues (4) words in sentences and (5) paired associates. A full discussion of these parts can be found in Dr. Carroll's paper.

The central question was that of the desirability of having a foreign language aptitude test for the Southeast Asian countries. With the exception of Laos and

Vietnam (for which a quantitatively pressing need did not appear to exist) and Indonesia and Philippines which were unrepresented in our workshop, all felt that there was indeed a real need for developing a foreign language aptitude test for language learners in their country.

Given desirability, the next most important questions were: (1) what languages would be involved? (2) would the test be marketable? i.e., what problems would be involved in gaining acceptance for the test? (3) how would it be used? (4) who would develop it and, (5) how would it be circulated? Since these are difficult, and in some cases sensitive questions, not all of them were resolved for each country. I will give you the conclusions that were arrived at country by country.

Singapore: Workshop members from Singapore felt that at least a Chinese and a Malay language version would be useful. The third major language group — Tamil and other Indian languages — was felt to be too small and complex a group to provide for at this time. The Singapore participants felt that Chinese and Malay language versions would be useful in screening primary school students for the English-medium schools and for diagnostic purposes to help determine the type of English program at the beginning of secondary school. Other uses might be made of the test to screen or place candidates in government or vocational training programs where some intensive English language instruction is required.

Malaysia: The workshop members from Malaysia felt that there was enough common ground linguistically, and in the intended use to which the test would be put, to consider a joint Singapore-Malaysia effort at developing Chinese and Malay language versions of the EMLAT or MLAT and that this research could perhaps be centered in Singapore but with collaboration from testing experts from Malaysia. The Research Unit of the Ministry of Education of Singapore was considered as a likely location for coordinating this research. It was felt desirable to plan the construction of a test for a level between the EMLAT and MLAT tests since a major testing population would be pre-secondary school age. Some Singapore and Malaysian participants felt that any efforts at screening students based on aptitude would be met with resistance (particularly from parents) and a good deal of data on test validity would have to be accumulated. Others felt that there would not be much of a problem in gaining acceptance for the test.

Thailand: The Thai sub-group of the workshop sessions also felt a definite need for a foreign language aptitude test which could be used to help determine when language training could most profitably be begun in the Thai school system. Such a test could also be administered at the MS 5 level (upper secondary) to determine — possibly screen — the input into some types of higher education (English majors at the universities, higher certificate teacher training programs, etc.) and for use generally as a diagnostic instrument.

It was learned that the DLI (Defense Language Institute) of the United States has a foreign language aptitude test for Thais but that this test had apparently not been used in Thailand. Information concerning other types of aptitude tests in the Thai language developed by the testing unit at Prasarnmit Teacher's College as well as those of the Thai Civil Service Commission was discussed.

It was felt that a Thai language version of the MLAT should definitely be attempted and that the University Development Commission English Language Centre might be the best place to coordinate the research since efforts at this point are quite diverse and potential researchers spread throughout the country.

Several problems concerning the actual test construction rather than validation techniques were discussed. Four problem areas were identified and while it was felt that these were common to all potential Southeast Asian language adaptations, they were particularly acute for Chinese. The major problem areas were: (1) lack of knowledge or inadequate knowledge of the writing system of the first language. This was felt to be the case for Chinese at nearly every level and for the other languages at pre-school and elementary school level. (2) Phonetic transcription — a new basis would have to be devised. (3) The number system would present problems in the number learning portion of an MLAT adaptation since there are some number systems which are too obvious. (4) The suitability of using a nonsense-language — such as is found in paired associates. (5) The difficult task of selecting appropriate syntactic problems for the "words in sentences" section of the adaptation.

None of these problems were resolved, but it was felt that the major problems were at least identified for potential adapters.

I think I've touch on the major points briefly — with the exception of all the detailed discussion about actual test construction for Thai, Chinese or Malay. On this matter, there are problems which would have to be worked out for some of the sub-tests of the MLAT. No version for use in Southeast Asia could be a mere translation. There will have to be a good deal of intuition, imagination and creativity in addition to hard, sustained work in order to develop reliable and valid aptitude test instruments for our eventual use.

WORKSHOP 8 — TECHNIQUES OF TEST CONSTRUCTION

ALAN DAVIES

The objective of this workshop was to consider techniques of testing and to relate these considerations not to standardized tests but to the teaching situation, to ways in which the techniques can be employed in the classroom situation.

The activities of the workshop proceeded through nine steps:

(1) We examined the general notion of test validity, and various types of validity required or desirable in a test.

(2) We examined quite a large collection of specimens of various test construction techniques useful for testing specific language skills or combinations of skills.

(3) All the group took an English test that used various testing techniques, the results of which were to serve as practical data for later discussion.

(4) The group divided into seven sub-groups, each of which set about constructing a test. Four groups constructed tests on reading, two on listening, one on speaking.

(5) We went through an item analysis of the test which the group had taken previously, to see how validity can be judged by item analysis

(6) The tests which had been constructed by the sub-groups were administered to the whole group, and scored.

(7) Each sub-group did an item analysis of the test which they had constructed.

(8) The whole group examined the correlations between the tests which had been constructed and administered.

(9) There was, finally, a general discussion of problems and difficulties encountered.

This quite practical approach obviously coincided with the desires of the workshop participants.

Some observations made in the course of the group's discussion are worth summarizing here.

Language teaching objectives are hard to state, and once they are stated, it is still hard to define exactly what a stated objective may *mean*. A test, however, says that *this* (i.e. a test item) is what is meant; i.e. in constructing tests we are at the same time specifying objectives. The two cannot be separated.

Testing will not do everything in language. In fact, it would seem that the only thing that can really be tested well is reception. Production *must* be taught, however, whether testable or not.

But there is a problem here — can something which can't be tested, be taught? Can we effectively teach something we can't really specify so that it is testable? Or are language teachers obliged to move amid generalities, like Arnold lecturing at Rugby on "moral character."

The basic problem, of course, is not in teaching or in testing, but in our knowledge of language itself. So long as we cannot specify what it is the native speaker knows, then we have a difficult job.

WORKSHOP 9 — ADMINISTRATIVE PROBLEMS OF ENGLISH LANGUAGE TESTING

MURAD BIN MOHD. NOR

The group thought it necessary first of all to distinguish what are the administrative problems which are peculiar to English language testing and what are the problems which are common to every kind of testing irrespective of the subject. It was also observed that, depending on the type and nature of the test, different problems may arise. In the discussion, therefore, a distinction was made between classroom tests which are administered in the course of the lesson, and national or regional tests which aim at finding out the standard and achievement of the language skill and are administered by the central or regional authorities.

In English language testing, the group chose three major areas of testing for evaluation:

(1) Oral production.

(2) Listening and Comprehension.

(3) Written tests.

Whether the test is to be conducted at the national or school level we may encounter more or less the same problems. At the national level the problems are as follows:

(1) Determination of standards.

(2) Coordination of marking by various examiners.

Besides these, the aim of the teaching of oral English, which varies from country to country, poses another problem in administering the tests.

In order to solve these problems it may be necessary to examine the method of teaching itself. In cases where English is taught as a subject and where the oral skill is not the main objective, the formal oral test at the national level may be done away with. It may be sufficient for a class teacher to assess his pupils' ability in oral English throughout the course of his teaching and to grade him accordingly if grading is required.

At any rate it was felt that one test at the end of the year or school career to determine oral skill may not be valid since various factors tend to influence one's oral performance. Hence the reliability of oral testing if it is only evidenced once during the pupil's school career, may be highly suspect. The best person to examine the oral ability therefore seems to be the classroom teacher himself even at the national level.

The administrative problems with regard to oral testing therefore are not so much in the conduct of the test itself but rather in the nature of the test which is so subjective that to design and administer a valid test poses difficulities which may be insurmountable. Since the result of the test cannot be truly objective perhaps an informal testing conducted by the class teacher himself is adequate if the aim is only to have a rough guide to pupils' abilities in oral production.

In testing listening comprehension skill the problems are somewhat similar to those of oral testing. Besides the difficulties in coding and grading the skill objectively we also have to take into consideration the wide range of teachers' abilities in spoken English. A candidate's performance may be to a certain extent influenced by the examiners' oral abilities We have yet to determine what kind of spoken English should be the model upon which we should base our pupils' ratings in listening and comprehension.

Perhaps a realistic standard of teacher abilities may be defined, and teachers of English should be trained to achieve this standard and examiners should be chosen only from those who have achieved the required standard.

In considering written tests we tried to divide it into three broad headings, namely: essay or composition, comprehension, grammar. In essay writing the main problem seems to be the lack of reliability or validity of the test. A free composition seems to be too subjective and some countries feel that this skill is too advanced within the context of their present-day curriculum. These countries suggest that a test involving sentence construction may be adequate.

Controlled essay too has been suggested to overcome this problem. In essay writing, whether one should test the language only or the ideas and content as well poses another problem. To overcome this it is suggested that both free and controlled essays should be used as a method of testing.

For comprehension and grammar it seems that a test in the objective form has considerable merit. This of course leads to many administrative problems. Construction of the objective type questions may be somewhat beyond the range of the average teacher's abilities if it is to be constructed and administered at the school level. At the national level, however, though possible, it may necessitate special training for a group of teachers in the methods of test construction. This obviously needs an administrative machinery within the educational system to carry it out.

Security of the test materials and the method of conducting the test itself, may require a well-planned, complicated and expensive testing organization. Some countries which have a decentralized system of education may find that this way does not fit well into their educational administrative structure.

Other administrative problems involved in the conduct of tests depend on two factors:

(*a*) Whether the conduct of the test is to be centralized or decentralized.

(*b*) The set-up of the educational administrative system of the country.

These two factors, however, influence not only the English language testing but all testing irrespective of subject.

Problems of centrally controlled testing are many. Among them are:

(1) Standardization.

(2) Coordination of examiners.

(3) Communication.

(4) Reliability and validity on a national scale.

Beside the above, teaching in such a system may tend to be inflexible, too conscious of syllabus and examination. It is felt that in spite of the above problems a fair amount of national standard or norm can be measured.

If on the other hand testing is conducted on a school or district basis the above problems, though they may still be present, are of more manageable magnitude. In this case the national standard may still be determined by closer supervision and adequate inspection of the schools.

The workshop therefore noted that both systems have considerable merit and both pose administrative problems in conducting tests in whatever subject.

Our workshop does not seem to give any specific recommendations on how to solve the problems, but what we have tried to do was to identify the problems. We hope in focussing our attention on the problems, each country may within its own educational administrative system work out the solutions to these problems.

WORKSHOP 10 — PREPARING AN ENGLISH PROFICIENCY TEST FOR ENGLISH TEACHER APPLICANTS IN SOUTH EAST ASIA

JOE D. PALMER

This workshop proceeded much like a graduate course in which the participants surveyed the field of testing and the factors involved in carrying out the kind of testing programs needed.

The sessions consisted of lectures and discussions.

The business of the workshop was the following:

1. **Comparisons of aptitude, proficiency and achievement tests, their places in school programs, and specifically the needs for English language proficiency tests in South East Asian countries.** The needs vary from urgent to non-existent.

2. **Analysis of the TOEFL, the MICHIGAN, and the MLA Advanced Proficiency Test**

These tests have among them these seven categories:

1. Grammar.
2. Vocabulary.
3. Listening.
4. Reading.
5. Writing.
6. Speaking.
7. Teaching.

It was found that TOEFL and MICH exclude speaking and teaching (above). The MLA excludes grammar and vocabulary as such. Categories 1 through 5 seem best suited to low-level proficiency testing. Categories 3 through 7 seem best suited to high-level proficiency testing. Categories 1 and 2 are most artificial, less related to real language. Speaking ability correlates with listening ability, and writing ability correlates with reading ability. Therefore, we *may* eliminate speaking and writing tests.

New writing tests comprise items based on grammar (structure) and lexicon. Sensitivity to style, tone, and register indicates a correlation to writing skill on the TOEFL (four 20-minute themes scored by professional readers correlate .72 with the TOEFL writing-section score).

Categories 6 and 7 are the most difficult physically to test.

3. **Mechanics of Testing.** Dr. Nababan (Indonesia) read a paper on some problems of validity in language testing. Examples of two pre-tests were studied in detail, particularly item writing and its pitfalls, and the need for linguistic sensitivity and an artistic grasp of language in finding good discriminators. The reasons for multiple-choice objective items were questioned. How to do item analysis was covered, specifically how to determine difficulty and discrimination indices. The minimum requirements of statistical validity (100 papers) 27 per cent division of top and bottom; 30 as minimum satisfactory discrimination index

were offered. Correlation among sections of the test was discussed. Administering tests and finding correlations between test-scores and academic success or later performance were discussed.

4. **Some areas of language study which have not been often used in writing items:** Research into rhetorical and critical insights and devices as the bases of testing linguistic sensitivity, particularly on high-level proficiency tests is needed. Examples were presented for consideration: lexical, grammatical, phonological, graphological, semantic and dialect deviation, register mixing, historical varieties, etymology, lexical deviation, metaphor, synechdoche, metonymy, neologism, parallelism, versification, expository, narrative, and argumentative writing, semantic oddities, symbolism, allegory, hyperbole, litotes, irony, innuendo, rhetorical question, apostrophe, ambiguity, homonymy, polysemy, etc.

5. **Some ideas for test-items:** In order to make the test a real language event ideas were sought, for example: naming poems, and choosing an accurate paraphrase.

6. **Conclusion:** Testing might move away from categories based on linguistic analysis to more traditional ways of regarding language — one which we might call a literary-philological way of analyzing language — to find items which more realistically test language ability than currently available proficiency tests do.

FINAL SESSION

Address by KHUN RUNJUAN INTARAKAMHANG *on behalf of the Ministry of Education of Thailand:*

Mr. Director, Mme. Chairman, Ladies and Gentlemen

As this regional seminar is closing, there are a few things I would like to say on behalf of the Ministry of Education of Thailand. First, I am happy to note that this is the first time that Thailand has received such a learned gathering in this area of specialization. Some of the speakers have come from as far as Nepal, England, Germany and the United States. We in Thailand feel that it is a great pleasure to have had you with us during the past seven days, and very much hope that your brief visit to our city has been not only beneficial but pleasant and enjoyable as well. For those of you who are not in a hurry to leave us, may I invite you to stay on for a few days, so that you will have time to know more of our city.

Secondly, I would like to observe that this seminar has aroused a great deal of enthusiasm and interest among our English teachers and educationists. In spite of the fact that schools and colleges here have already opened, it is still noticeable that Thai participants outnumber the participants from overseas. This enthusiastic response from our professional people gives us good reason to believe that there will be development in the methods of learning English in this country and that in fact many better approaches will be adopted.

Finally, I should like to express on behalf of the Bangkok organizing committee, words of admiration and appreciation to Mrs. Tai Yu-lin, Director of the Regional English Language Centre, for her inspiring leadership and initiative in organizing this seminar.

I also record words of thanks to Mrs. Mayuri Sukwiwat, Director of the English Language Centre of the University Development Commission here in Thailand and her entire staff for their hard work before and all through the seminar. My thanks to the various international agencies in Bangkok for their assistance and cooperation. And if I may, I should like also to include my personal appreciation and thanks to my colleagues.

And before I end, may I congratulate all the participants for having contributed to the success of this seminar.

FINAL SESSION

Address by TAI YU-LIN *Director of the Regional English Language Centre.*

Mr. Director, Distinguished Guests, Ladies and Gentlemen:

Before I declare this seminar closed, I would like to exercise my privilege as Director of the RELC to say a few words. In a short period of five and a half days, we have covered a great variety of areas relevant to English Language Testing. We have enjoyed the scholarly papers presented by eminent speakers from within and outside of South East Asia, on topics ranging from general principles of language testing, through the philosophy and the challenges ahead, specifically relevant to South East Asia. We have attempted to examine as thoroughly as has been feasible, in a short time, the exciting developments in recent years in the field of language testing and their potentialities in the South East Asian context. Our discussions have been open and frank; we have made realistic appraisals of our own situations in an attempt to identify present problems and to search for suitable solutions. The seminar has been constructive and rewarding, not only because of the new information we have been able to obtain but also because of the hard thinking it has stimulated. It has been a meeting of minds between leaders of education in South East Asia who have to think in terms of what is workable within the context of their educational systems, and today's leaders in language testing from universities, research centres and testing services of four continents. The information given to both sides is surely of immense value and promises hope of more adequate testing instruments to be developed for second and foreign language teaching in the future.

There has been an earnest and sincere attempt at international cooperation to insure progress and advancement in education and learning. There have also been definite signs that the six-day session has helped to promote closer and greater understanding and good will between educationists in South East Asia and the world. That the discussions have been lively and that the workshops have produced some interesting and fruitful findings and results, has been much due to the friendly and cooperative spirit that has continued throughout the seminar and the efficient manner in which the chairmen, the rapporteurs, and the many reporting officers have conducted the meetings. On behalf of the Regional English Language Centre, I would like to express to one and all my profound appreciation.

This expression of appreciation would not be complete without mentioning the wholehearted cooperation and generosity of the Ministry of Education of Thailand. We are much indebted to the dedicated and hardworking organizing committee under the able leadership of Khun Runjuan Intarakamhang, M. L. Boonlua Debyasuvarn, and Mrs. Mayuri Sukwiwat, for the excellent manner in which this seminar has been organized. I would also like to mention specifically the capable leadership in various areas of Mrs. Vipa Chulajata and Miss Tongbai Chittamongkol. Special mention must be made of the most efficient secretariat, which has processed over 175,000 sheets of text material and over 88 million words for the seminar.

It is understood that delegations of several member countries are already making concrete plans of organizing national seminars on regional or district basis immediately on their return to their respective countries. They will use the materials of this seminar to disseminate information and to generate new thinking among educationists in their respective countries in order that your deliberations and the results of this seminar have far-reaching effects in South East Asia.

We are deeply grateful to His Excellency the Minister of Education, Thailand, for the personal attention he has given to the seminar and the warm hospitality he has shown.

Those of us who are leaving Bangkok at the end of the seminar will take back with us fond memories of our stay in this beautiful city where the rich cultural heritage speaks for itself. The beautiful cultural show which we enjoyed on the evening of the sumptuous dinner given by His Excellency the Minister of Education will long be remembered by us all.

We also owe thanks for the graciousness and generosity of the staff members of the University of Pittsburgh Project in Thailand, the British Embassy, the British Council, the American Universities Alumni Language Centre, and many others who have made our evenings in the past five days so pleasant and relaxing after the long sessions during the day.

So many people rendered assistance in one form or another, in organizing this seminar and in making our stay in Bangkok memorable. I wish to thank all our friends and helpers for the services and courtesies extended. I also wish to thank the publishers and educational agencies for their cooperation in setting up the fine display of books and other exhibits to add to the success of the seminar.

On behalf of the Regional English Language Centre I wish our distinguished delegates and participants who have done us the honor of coming from afar to attend this seminar a very pleasant journey home. We hope you have found the week spent with us in South East Asia profitable, and we look forward to having the pleasure of welcoming you to another RELC occasion in future. Thank you all again for your cooperation and good will.

I now officially declare this seminar closed.

SEMINAR OFFICIALS

Seminar Director: Tai Yu-Lin
Director, Regional English Language Centre

Seminar Administrative Officer: Norman de Souza
Registrar and Bursar, RELC

Seminar Planning Committee:

William B. Owen, Chairman
  Specialist in Evaluation and Testing, RELC

Alan D. Moller
  Specialist in Instructional Materials, RELC
  Coordinator of Workshop Activities

Kenneth W. Bale
  Specialist in A/V Aids, RELC
  Documents Officer

Roy Chérrier
  Specialist in Linguistics, RELC
  Editor, Seminar Report

Yolanda Beh
  Librarian and Head of Information Centre, RELC
  Exhibition Supervisor

Seminar Local Organizing Committee

Khun Runjuan Intarakamhang, Chairman
  Director, External Relations Division
  Ministry of Education

M. L. Boonlua Debyasuvarn, Adviser
  Member, RELC Coordinating Committee
  Acting Dean of Liberal Arts
  Silpakorn University

Mayuri Sukwiwat, Secretary
  Director, English Language Centre
  University Development Commission

Vipa Chulajata, Assistant Secretary
  Officer-in-charge of SEAMEO Affairs
  External Relations Division
  Ministry of Education

Choob Buspavanich
  External Relations Division
  Ministry of Education

Tongbai Chittamongkol
  Head, English Department
  Bangkok Commercial College

Wit Thapthiang
  Librarian, College of Education

Napa Bhongbhibhat
  Educational Information Division
  Ministry of Education

P. B. Prabha
  College of Education

Panee Boonthavi
  Faculty of Arts
  Chulalongkorn University

Urai Bhundhoombhoad
  College of Commerce

Malinee Chandavimol
  Demonstration School
  Chulalongkorn University

Nanda Kowongse
  Thonburi Technical College

Savitri Suwansathit
  External Relations Division
  Ministry of Education

Siraporn Sangprabha
  Department of Teacher Training
  Ministry of Education

Saowanee Indrabhakti
  College of Education

Pranee Dhanachanandha
  College of Education

SPEAKERS AND WORKSHOP LEADERS

Ralph P. Barrett
Head, University of Pittsburgh English Project
Thammasat University
Bangkok, Thailand

Eric W. Bauer
Director of Language Services
Monash University
Clayton, Victoria
Australia

Duong Thanh Binh
Professor of English
Faculty of Pedagogy
University of Saigon
Saigon, Vietnam

Eugene J. Briere
Associate Professor of Linguistics
University of Southern California
Los Angeles, California,
U.S.A.

John B. Carroll
Senior Research Psychologist
Educational Testing Service
Princeton, New Jersey
U.S.A.

Alan Davies
Head, Department of English
Tribhuvan University College
Kirtipur, Nepal

Robert B. Kaplan
Director, English Communication Program for Foreign Students
University of Southern California
Los Angeles, California
U.S.A.

Clodoaldo H. L[illegible]ocadio
Acting Assistant Director of Public Schools
Bureau of Public Schools
Manila, Philippines

Lim Kiat Boey
Senior Lecturer
Faculty of Education
University of Malaya
Kuala Lumpur, Malaysia

M. Muchnilabib
Head of Language Laboratory
English Department
IKIP Malang
Malang, Indonesia

Murad bin Mohd. Nor
Controller of Examinations
Ministry of Education
Kuala Lumpur, Malaysia

P. W. J. Nababan
Senior Lecturer in Linguistics and English
IKIP Malang
Malang, Indonesia

Ng Kuen Seong
Specialist Inspector/English as a Second Language
Ministry of Education
Singapore

Gerhard Nickel
Lehrstuhl Anglistik
Abt. Anglistik
Universitaet Stuttgart
Stuttgart, Germany

Richard B. Noss
Chief Adviser, English Language Center
University Development Commission
Bangkok, Thailand

Joseph Palmer
Rockefeller Foundation
Bangkok, Thailand

Oliver Seet Beng Hean
Lecturer, English Studies Department
Teachers' Training College
Singapore

Charles R. Sheehan
Regional Language Supervisor
U.S. Embassy
Bangkok

Bonifacio P. Sibayan
Dean, Graduate School
Philippine Normal College
Manila, Philippines

Pinhkham Simmalavong
Director of Studies
Ecole Supérieure de Pédagogie
Vientiane, Laos

Brian Deane Smith
Advisor in Testing, English Language Center
University Development Commission
Bangkok, Thailand

Huynh Dinh Te
Associate Professor of English and Linguistics
Faculty of Letters
University of Hue
Hue, Vietnam

Banluc Tinpanga
Head, Department of Humanities and Social Science
The College of Education
Songkhla, Thailand

Pramono Tirtopramono
Secretary, English Coordinating Board
Ministry of Education
Djakarta, Indonesia

John A. Upshur
Director, Testing and Certification Division
English Language Institute
University of Michigan
Ann Arbor, Michigan
U.S.A.

LIST OF SEMINAR PARTICIPANTS

1. SEAMES Officials

General Netr Khemayodhin
Director, SEAMES

Soemantri Hardjoprakoso
Assistant Director I

Khoo Eng Choon
Assistant Director II

Mehdi Dulyachinda
Fund Raising Consultant

John Davis Williams
Ford Foundation Consultant to SEAMES

Thavisakdi Srimuang
Program Development Assistant

Marcos S. Ramos
Program Analyst

2. Official Delegates from SEAMEO Countries

Indonesia:

P. Tirtopramono,
Head of Delegation
Secretary, English Coordinating Board
Ministry of Education
Djakarta, Indonesia

Samsuri
Member, RELC Coordinating Committee
Rector, IKIP Malang
Malang, Indonesia

M. Muchnilabib
Head of Language Laboratory
English Department
IKIP Malang
Malang, Indonesia

P. W. J. Nababan
Senior Lecturer in Linguistics and English
IKIP Malang
Malang, Indonesia

Laos:

Pinhkham Simmalavong,
Head of Delegation
Member, RELC Coordinating Committee
Director of Studies
Ecole Supérieure de Pédagogie
Vientiane, Laos

Chanthouma Souvannakhily
Ecole Supérieure de Pédagogie
Vientiane, Laos

Somsack Saythongphet
Principal, Fa Ngum Comprehensive High School
Vientiane, Laos

Tou Fu
Principal, College of Samthong
Xiengkhouang, Laos

Malaysia:

Murad bin Mohd. Nor,
Head of Delegation
Director of Higher Education
Educational Planning and Research Department
Ministry of Education
Kuala Lumpur, Malaysia

Yap Hong Kuan
Member, RELC Coordinating Committee
Chief Education Officer, Selangor
Kuala Lumpur, Malaysia

Bernadette Koay
Examination Specialist
Examination Syndicate
Ministry of Education
Kuala Lumpur, Malaysia

Lim Kiat Boey
Senior Lecturer
Faculty of Education
University of Malaya
Kuala Lumpur, Malaysia

Philippines:

Jose T. Enriquez,
Head of Delegation
Member, RELC Coordinating Committee
Acting Director of Public Schools
Bureau of Public Schools
Manila, Philippines

Clodoaldo H. Leocadio
Acting Assistant Director of Public Schools
Bureau of Public Schools
Manila, Philippines

Bonifacio P. Sibayan
Dean, Graduate School
Philippine Normal College
Manila, Philippines

Melba Tugade
Supervising Educational Specialist
Research, Evaluation, and Guidance Division
Bureau of Public Schools
Manila, Philippines

Fe Manza *
Chief, Elementary Education Division
Bureau of Public Schools
Manila, Philippines

Inez Q. David *
Consultant, Elementary Education Division
Bureau of Public Schools
Manila, Philippines

Singapore:

Molly Ng Kuen Seong,
Head of Delegation
Specialist Inspector/English as a Second Language
Ministry of Education
Singapore

Low Kee Cheok
Member, RELC Coordinating Committee
Specialist Inspector/English Language and Literature
Ministry of Education
Singapore

Oliver Seet Beng Hean
Lecturer, English Studies Department
Teachers' Training College
Singapore

Lee Sow Ling
Research Officer
Research Unit
Ministry of Education
Singapore

Lieut. Thomas Tan†
Language Training Department
Ministry of Interior and Defence
Singapore

Lieut. Philip Lee†
Language Training Department
Ministry of Interior and Defence
Singapore

Thailand:

Prachoom Dabbasuta,
Head of Delegation
Head, Department of Languages
Katsesart University
Bangkok, Thailand

Kanda Sittachitta
Head, Department of Modern Languages
Chiengmai University
Chiengmai, Thailand

Wanida Amundson
Chulalongkorn University
Bangkok, Thailand

Komkai Chongcharoensuk
Dept. of Teacher Training
Ministry of Education
Bangkok, Thailand

Prayad Laksana-ngam
Dept. of Teacher Training
Ministry of Education
Bangkok, Thailand

Banlue Tinpanga
Head, Dept. of Humanities and Social Science
College of Education
Songkhla, Thailand

***Participation financed by Government of Philippines.**
† Participation financed by Government of Singapore.

Iadsiri Liengjayetz
Dept. of Secondary Education
Ministry of Education
Bangkok, Thailand

Thomas A. Durr
Dept. of Secondary Education
Ministry of Education
Bangkok, Thailand

Chomtawin Niemtan
Dept. of Elementary Education
Ministry of Education
Bangkok, Thailand

Siriporn Limtrakarn
Dept. of Vocational Education
Ministry of Education
Bangkok, Thailand

Chamnian Srikrajang
Dept. of Vocational Education
Ministry of Education
Bangkok, Thailand

Vietnam:

Duong Thanh Binh,
Head of Delegation
Member, RELC Coordinating Committee
Professor of English
University of Saigon
Saigon, Vietnam

Huynh Dinh Te
Professor of Linguistics
University of Hué
Hué, Vietnam

Le Van
Professor of English
University of Saigon
Saigon, Vietnam

Tran Trong Hai
Inspector, Specialist in English Testing
Ministry of Education
Saigon, Vietnam

3. Official Representatives from non-SEAMEO Governments

Rytsa Finkelstein
Canadian Embassy
Bangkok, Thailand

J. P. Stocks
Asst. Chief Inspector of Schools
Education Department
Hong Kong

M. N. Cardiff
British Council Representative
Bangkok, Thailand

Jack Juergens
United States Information Service
U. S. Embassy
Bangkok, Thailand

Ramesh Mohan
Central Institute of English
Hyderabad, India

4. Other Participants

Joanne Ackerman, AUA Language Center, Bangkok

Paul L. Aiken, Ford Foundation, National Institute of Development Administration, Bangkok

Siwalee Ampai, Dept. of Vocational Education, Ministry of Education, Bangkok

Amari Amaranonta, Silpakorn University, Bangkok

Orasa Angsusingh, Dept. of Vocational Education, Ministry of Education, Bangkok

Sdee Ansusinha, Dept. of Secondary Education, Ministry of Education, Bangkok

A-gnum Arsawes, Dept. of Vocational Education, Ministry of Education, Bangkok

Supeewong Attasart, Dept. of Teacher Training, Ministry of Education, Bangkok

Lyle Bachman, English Language Center, University Development Commission, Bangkok

Nipon Banjong, Dept. of Teacher Training, Ministry of Education, Bangkok

Ralph P. Barrett, University of Pittsburgh English Project, Thammasat University, Bangkok

D. Bell, British Council, Bangkok

Virginia B. Berk, Military Academy, Bangkok

Urai Bhundhoombhoad, Dept. of Vocational Education, Ministry of Education, Bangkok

Suwaree Bijaisoradat, Dept. of Teacher Training, Ministry of Education, Bangkok

Duangduan Bisalputra, Chulalongkorn University, Bangkok

Rae F. Blessing, c/o Australian Embassy, Vientiane, Laos

Brian Blomfield, Chung Chi College, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Jackie Boehme, AUA Language Center, Bangkok

Smaiswat Bongsadadi, Silpakorn University, Nakorn Pathom, Thailand

Tienmanee Boonjoon, Dept. of Teacher Training, Ministry of Education, Bangkok

Suchitra Boonsa-ard, Dept. of Teacher Training, Ministry of Education, Bangkok

Thida Boontham, Dept. of Secondary Education, Ministry of Education, Bangkok

P. B. Botsman, Papua and New Guinea Institute of Technology, Lae, New Guinea

Jeanne M. Brand, English Language Center, Asian Institute of Technology, Bangkok

Eugene Briere, University of Southern California, Los Angeles, Calif., U.S.A.

Bert H. Brewer, Military Academy, Bangkok

Geoff Burke, Ecole Supérieure de Pédagogie, Vientiane, Laos

M. Broughton, British Council, Bangkok

J. Marvin Brown, AUA Language Center, Bangkok

Nikom Buddhamatya, Dept. of Teacher Training, Ministry of Education, Bangkok

Joan A. Bugbee, Dept. of Teacher Training, Ministry of Education, Bangkok

Apron Bukhamana, Chulalongkorn University, Bangkok

Charoontasana Bukhamana, Chulalongkorn University, Bangkok

Nonglak Bunnag, Dept. of Teacher Training, Ministry of Education, Bangkok

Samang Burtintavanij, Dept. of Vocational Education, Ministry of Education, Bangkok

Sukhuman Butranon, Dept. of Teacher Training, Ministry of Education, Bangkok

John B. Carroll, Education Testing Service, Princeton, N.J.

Peter Cavaye, British Council, Bangkok

Sirikul Chairatana, Civil Service Commission, Bangkok

Panta Chaitiamvong, Chulalongkorn University, Bangkok

Somkid Chaiyunboon, Chulalongkorn University, Bangkok

Chuthai Chalamat, Dept. of Teacher Training, Ministry of Education, Bangkok

Malinee Chandavimol, Demonstration School, Chulalongkorn University, Bangkok

Paka Chandvirach, Royal Thai Navy School, Bangkok

Thanomsri Changtrakul, Dept. of Vocational Education, Bangkok
Niramol Changwatchai, Chulalongkorn University, Bangkok
Varunee Chantaranuvat, Chulalongkorn University, Bangkok
Anong Charanacupta, Dept. of Teacher Training, Ministry of Education, Bangkok
Nanthana Charoenpot, Dept. of Vocational Education, Ministry of Education, Bangkok
Ahporn Charriburus, Chulalongkorn University, Bangkok
Krongtong Chavalitdhamrong, National Institute of Development and Administration, Bangkok
Saroj Chavanaviraj, Ministry of Foreign Affairs, Bangkok
Prasong Chingchai, Supreme Command Headquarters, Bangkok
Ratjanee Chirtkeatisak, Civil Service Commission, Bangkok
Yubol Cholaseuks, Dept. of Vocational Education, Bangkok
Patana Chuchua, Dept. of Vocational Education, Bangkok
Sophia Chulajata, Dept. of Vocational Education, Bangkok
Prasit Chumpa, Dept. of Teacher Training, Bangkok
D. Churchill, Chulalongkorn University, Bangkok
David Jeffery Cobb, Asian Institute of Technology, Bangkok
Conchita M. Cruz, Central Luzon State University, Munoz, Nueva Ecija A-632, Philippines
Alan Davies, c/o British Council, Kathmandu, Nepal
John N. Dent-Young, Chiengmai University, Chiengmai, Thailand
Bhanor Devahastin Na Ayudhya, Thai Army Field Forces, Bangkok
Chomphit Devajin, Chulalongkorn University, Bangkok
Suraja Devakul, Armed Forces Academic Preparatory School, Bangkok
Dhalad Dhiensiri, Dept. of Vocational Education, Bangkok
Praopan Duhnma, Dept. of Teacher Training, Bangkok
Wilai Disathien, Dept. of Vocational Education, Bangkok
A. C. Dunlop, British Council, Bangkok
Suraphoe Eenyam, Dept. of Teacher Training, Bangkok
Colin R. Elliott, Chulalongkorn University, Bangkok
David Eskey, University of Pittsburgh English Project, Thammasat University, Bangkok
Timothy W. Flegel, Chulalongkorn University, Bangkok
Boonao Gasisilpa, Dept. of Vocational Education, Bangkok
Nantha Gohwong, Thonburi Technical Institute, Thonburi, Thailand
Dorris Gold, Fulbright Foundation, Bangkok
James A. Golman, U.S. Operations Mission, Bangkok
Vilas Hannarkin, Dept. of Teacher Training, Bangkok
A. Harkess, Silpakorn University, Nakorn Pathom, Thailand

Jimmy G. Harris, English Language Center, University Development Commission, Bangkok

Hiroyoshi Hatori, Tanashi-shi, Tokyo 188, Japan

R. Hawkey, British Council, Bangkok

Sukon Hiranburana, Dept. of Vocational Education, Bangkok

Marjorie Hoath, AUA Language Center, Bangkok

Suporn Homchan, Dept. of Vocational Education, Bangkok

Susaki Hongsakul, Dept. of Vocational Education, Bangkok

Marilyn Hulbert, Lembaga Indonesia-America, Djakarta, Indonesia

Areerut Hurunro, Dept. of Vocational Education, Bangkok

Nantana Husain, Royal Thai Air Forces School, Bangkok

Pianjit Inchan-Narong, Dept. of Vocational Education, Bangkok

Saowance Indrabhakti, College of Education, Bangkok

Prapatsorn Intarakamhaeng, Dept. of Vocational Education, Bangkok

Kumpol Isarangkura, Chulalongkorn University, Bangkok

Antonio J. Isarankura, Thammasat University, Bangkok

Chomsri Israngura, Chulalongkorn University, Bangkok

Peter Isles, Department of Technical and Economic Cooperation, Bangkok

Boonperm Ittirat, Dept. of Teacher Training, Bangkok

Manop Jamkrajang, Dept. of Teacher Training, Bangkok

Jarulak Jarupat, Dept. of Vocational Education, Bangkok

Keith Lynden Jennings, Dept. of Education and Service, Canberra, Australia

Permsuk Jermsawasdi, Chulalongkorn University, Bangkok

Barbara Jones, AUA Language Center, Bangkok

Pimonphan Kaehacharoen, Dept. of Secondary Education, Bangkok

Kay Kambhu, AUA Language Center, Bangkok

Wanna Kampiranon, Dept. of Secondary Education, Bangkok

Suwaphit Kanithayon, Dept. of Teacher Training, Bangkok

Robert B. Kaplan, University of Southern California, Los Angeles, Calif.

Wanida Kasipar, Dept. of Teacher Training, Bangkok

Somsak Kaysapakul, Silpakorn University, Nakorn Pathom, Thailand

Uraivan Kemprasit, Dept. of Vocational Education, Bangkok

Tuanchai Keowopas, Dept. of Elementary Education, Bangkok

Somsri Ketudat, Civil Service Commission, Bangkok

Sanit Khemtong, National Institute of Development Administration, Bangkok

Praguysee Khosongnoen, Dept. of Vocational Education, Bangkok

Orasa Khunawat, Dept. of Technical and Economic Cooperation, Bangkok

Supa Kitchathorn, Dept. of Teacher Training, Bangkok

Vanida Kitjalak, Dept. of Teacher Training, Bangkok

William Klausner, Asia Foundation, Bangkok

Hom Klayanonda, Dept. of Teacher Training, Bangkok

Theodore A. Klein, Language Center, Royal Thai Navy, Bangkok

Gerald E. Knight, American Institute of Southeast Asia, Bangkok

Jonathan Knowles, New Asia College, Chinese University of Hong Kong, Kowloon, Hong Kong

Aramsri Kompanthong, Dept. of Vocational Education, Bangkok

Montean Kongritrachan, Dept. of Vocational Education, Bangkok

Wasana Kovitaya, Chulalongkorn University, Bangkok

Ourora-Prachitra Krairiksh, Civil Service Commission, Bangkok

Nalinee Krikunasai, Dept. of Vocational Education, Bangkok

Foongfuang Kruatrachue, Pranakorn College of Education, Thailand

Robert E. Krug, American Institutes for Research, Bangkok

Anna Kwan-Terry, University of Singapore, Singapore

Suphab Kwanyuen, Dept. of Teacher Training, Bangkok

Tipawan Lachawanich, Chulalongkorn University, Bangkok

Sermsri Laksanasiri, Dept. of Elementary Education, Bangkok

Yaowarat Laosinchai, Dept. of Secondary Education, Bangkok

Yauwalak Leelachart, Dept. of Teacher Training, Bangkok

Daranee Leelaphatr, Dept. of Vocational Education, Bangkok

Lim Tong-Tang, Language Center, Nanyang University, Singapore

Rampai Limanond, Dept. of Vocational Education, Bangkok

M. Long, British Council, Bangkok

Graham J. Lucas, Asia Foundation, Bangkok

Jannes Albert Frans Lumbantobing, Universitas H.K.B.P. Nommensen, P. Siantar, Sumut, Indonesia

Robert Walker Lynn, Language Center, Nanyang University, Singapore

Max Maisel, U.S. Information Service, JUSPAO, American Embassy, Saigon, Vietnam

Krasae Malyaporn, Dept. of Teacher Training, Bangkok

Chao Maneewongse, Dept. of Teacher Training, Bangkok

Urai Mangtrisan, Dept. of Teacher Training, Bangkok

Prongpan Maninandana, Dept. of Vocational Education, Bangkok

Prowmat Manoratana, Dept of Teacher Training, Bangkok

B. Ellen Mathews, AUA Language Center, Bangkok

Jim McFetridge, Canadian University Services Overseas, Bangkok

Paul McGinnis, Canadian University Services Overseas, Bangkok

James D. McHale, U.S. Information Services, Singapore

Runjuan Meanapradit, Dept. of Vocational Education, Bangkok

Prayoon Metong, Dept. of Teacher Training, Bangkok

Suta Mettraipan, Dept. of Teacher Training, Bangkok

Jean-Marc Montguerre, SEATO, Bangkok

Nantana Moongvicha, Dept. of Vocational Education, Bangkok
M. Murphy, IKIP Negeri, Bandung, West Java, Indonesia
Shri Shaik Mowla, General Institute of English, Bangalore, India
Vipha Muttamara, Dept. of Vocational Education, Bangkok
Betty Jo Myers, AUA Language Center, Bangkok
Nida Narumitsuvimon, Dept. of Vocational Education, Bangkok
Gerhard Nickel, Abt. Linguistik, University of Stuttgart, Stuttgart, Germany
Richard B. Noss, English Language Center, University Development Center, Bangkok
Kamala Nutabandhu, Dept. of Vocational Education, Bangkok
Ladawan Nutsati, Dept. of Teacher Training, Bangkok
Yoshio Ogawa, Setagaya, Tokyo, Japan
Kenji Ohtomo, English Language Education Council, Tokyo, Japan
H. S. Otter, University of Cambridge Local Examinations Syndicate, Cambridge, England
M. O'Shea, British Council, Bangkok
Joseph Palmer, Univ. of Pittsburgh English Project, Thammasat University, Bangkok
Phensri Panitying, Dept. of Vocational Education, Bangkok
Sunan Panmeechao, Chulalongkorn University, Bangkok
Orapin Pensupha, Dept. of Vocational Education, Bangkok
Varan Patitamrak, Dept. of Vocational Education, Bangkok
Lyle Petch, Canadian University Services Overseas, Bangkok
Viroj Phanichkit, Chulalongkorn University, Bangkok
Walee Phantmanee, Chulalongkorn University, Bangkok
Sucha Na Phatalung, Dept. of Teacher Training, Bangkok
Vimolvan Phatharodom, Division of Educational Planning, Ministry of Education, Bangkok
Bradley Philips, Education Department, Kota Kinabalu, Sabah, Malaysia
Sumon Pholpark, Dept. of Secondary Education, Bangkok
Suphaphan Phuangchanpetch, Chulalongkorn University, Bangkok
Daranee Phummawon, Dept. of Teacher Training, Bangkok
Laeka Piya-Ajariya, Chulalongkorn University, Bangkok
Natsawong Pleumpunya, Dept. of Teacher Training, Bangkok
Sunit Prabhasawat, Dept. of Educational Techniques, Ministry of Education, Bangkok
Boonpriab Prahcanboribal, Dept. of Teacher Training, Bangkok
Pailin Pratoomsri, Dept. of Vocational Education, Bangkok
Banjapol Prommin, Royal Thai Air Forces School, Bangkok
Ua-a-ree Pongissawaranun, Dept. of Teacher Training, Bangkok
Surai Pongtongcharoen, Dept. of Teacher Training, Bangkok
Wandee Potprasart, Dept. of Vocational Education, Bangkok

Surarpon Puangmali, Chulalongkorn University, Bangkok
Poonna Pula, Dept. of Teacher Training, Bangkok
Vanida Pulsirivong, Chulalongkorn University, Bangkok
Arunee Puncharoen, Dept. of Vocational Education, Bangkok
Smai Punyasukh, Dept. of Vocational Education, Bangkok
Chongkol Purakam, Dept. of Vocational Education, Bangkok
Wirawatt Putawong, Dept. of Vocational Education, Bangkok
Daniel J. Rajah, Teachers' Training College, Singapore
Lucille M. Ramish, Language Center, Nanyang University, Singapore
Opas Ratanaburi, Supreme Command Headquarters, Bangkok
Bhaibulya Ratanamungala, Division of Educational Information, Ministry of Education, Bangkok
Puangchan Ratanaporn, Dept. of Vocational Education, Bangkok
S. Ratchichandran, MARA Institute of Technology, Petaling Jaya, Selangor, Malaysia
J. Ratnaike, UNESCO Regional Office, Bangkok
Erica Rawlinson, Chulalongkorn University, Bangkok
D. Robertshaw, British Council, Bangkok
Charlotte Reid, c/o British Council, Kuala Lumpur, Malaysia
Orasa Roongsang, Chulalongkorn, University, Bangkok
Pitthaya Rungratri, Dept. of Teacher Training, Bangkok
Orapan Saichoomin, Dept. of Vocational Education, Bangkok
Kanikar Saligupta, Silpakorn University, Nakorn Pathom, Thailand
Kritsir Samabuddhi, Civil Service Commission, Bangkok
Suta Sampawapon, Dept. of Vocational Education, Bangkok
John Samuel, Pengkalan Chepa, Kota Bharu, Kelantan, Malaysia
Duangduan Sangchai, Chulalongkorn University, Bangkok
Surapee Sangkapichai, Chulalongkorn University, Bangkok
Uraikarn Sansook, Dept. of Secondary Education, Bangkok
Khuanchai Sanunvanich, Dept. of Vocational Education, Bangkok
Panninee Sararik, Chulalongkorn University, Bangkok
Ampai Sarayudpitak, Dept. of Vocational Education, Bangkok
Pranom Sarojamarn, Thammasat University, Bangkok
Sunan Sastravaha, Dept. of Vocational Education, Bangkok
Niramon Satidtong, Chulalongkorn University, Bangkok
Pachara Sattabanasuk, Silpakorn University, Nakorn Pathom, Thailand
Paga Sattayatam, Chulalongkorn University, Bangkok
Neal Scott, Chulalongkorn University, Bangkok
M. A. Sculthorp, Language Centre, University of Kent, Canterbury, Kent, England
Paul A. Schwarz, American Institutes for Research, Bangkok
Pornpimol Senawong, Dept. of Vocational Education, Bangkok

Charles R. Sheehan, U. S. Embassy, Bangkok
Marion E. Siffermann, Military Academy, Bangkok
Praphasri Siha-umphai, Chulalongkorn University, Bangkok
Kaekai Simargool, Dept. of Secondary Education, Bangkok
Patama Singhavadhaka, Dept. of Vocational Education, Bangkok
Wilars Singhawisai, Civil Service Commission, Bangkok
Dara Siricharoen, Dept. of Teacher Training, Bangkok
Oo-sa Siriwong, Dept. of Vocational Education, Bangkok
Somsong Sitalayan, Dept. of Vocational Education, Bangkok
Somsook Smarnchai, Dept. of Vocational Education, Bangkok
Saiyud Smitalumpha, Chulalongkorn University, Bangkok
Brian Deane Smith, English Language Center, University Development Center, Bangkok
Michael Smithies, Univ. of Pittsburgh English Project, Thammasat University, Bangkok
Kamploy Snidvongs SEATO, Bangkok
Ruth Snyder, Thai Army Field Services, Bangkok
Ivy Soh, Teachers' Training College, Singapore
Rapee Somaphuti, Dept. of Vocational Education, Bangkok
Kalika Sombatsiri, Dept. of Vocational Education, Bangkok
Yoobha Songsiri, Dept. of Teacher Training, Bangkok
Somsri Sonthisupaya, Dept. of Vocational Education, Bangkok
Tepin Soontaros, Dept. of Teacher Training, Bangkok
Maroon Soopharb, Chulalongkorn University, Bangkok
H. T. Sparby, Canadian Embassy, Bangkok
Sanan Srijack, Dept. of Teacher Training, Bangkok
Aim-Orn Srinilta, Dept. of Vocational Education, Bangkok
Suwakit Sripathan, Dept. of Teacher Training, Bangkok
C. W. Stevens, Fa Ngum Comprehensive School, Vientiane, Laos
Prayuratna Stholasuta, Dept. of Secondary Education, Bangkok
Sukjai Suavansri, National Educational Council, Bangkok
Suthinee Suktrakul, Thammasat University, Bangkok
Narumol Sundaranu, Chulalongkorn University, Bangkok
Tassaniya Suprichakorn, Dept. of Teacher Training, Bangkok
Manit Suttajit, Uttaradit Teachers College, Uttaradit, Thailand
Chusri Suvannochot, Dept. of Teacher Training, Bangkok
Pinyalak Suwan, National Institute of Development Administration, Bangkok
Sumitra Suwanabha, Chiengmai University, Chiengmai, Thailand
Swasdi Suwanangsorn, Division of Educational Information, Ministry of Education, Bangkok
Savitri Suwansathit, Division of External Relations, Ministry of Education, Bangkok

Tasanee Suwatana, Dept. of Vocational Education, Bangkok
Sisamorn Tangsubuta, Military Academy, Bangkok
Aura Tanskul, Silpakorn University, Nakorn Pathom, Thailand
Prasert Tanskul, Dept. of Teacher Training, Bangkok
Prapai Tantrakool, Dept. of Secondary Education, Bangkok
Ellen Taylor, AUA Language Center, Bangkok
Mary Tay Wan Joo, Language Centre, Nanyang University, Singapore
Gail Taylor, Canadian University Services Overseas, Bangkok
Paibool Tevaragsa, Chulalongkorn University, Bangkok
Suwannee Teng-Amnuay, Chulalongkorn University, Bangkok
Saiswat Thammapimuk, Chulalongkorn University, Bangkok
Pranee Thanachanan, Dept. of Teacher Training, Bangkok
Sanit Thawisomboon, National Institute of Development Administration, Bangkok
William G. Thebus, Defense Language Institute, MACTHAI-JUSMAG, Bangkok
Niramol Thiradharm, Dept. of Vocational Education, Bangkok
Murray Thompson, Canadian University Services Overseas, Bangkok
Edwin Thumboo, University of Singapore, Singapore
Kannikar Tongprayoon, National Institute of Development Administration, Bangkok
Puapun Tongyode, Chulalongkorn University, Bangkok
Jonathan Trench, Asian Institute of Technology, Bangkok
Sinee Tubplang, Dept. of Teacher Training, Bangkok
Nisa Udomphol, Chulalongkorn University, Bangkok
Sayomporn Umbhol, Dept. of Vocational Education, Bangkok
John A. Upshur, English Language Institute, University of Michigan, Ann Arbor, Michigan
Pleuang Uraipongse, Dept. of Vocational Education, Bangkok
Ladaporn Vanitsthian, Thammasat University, Bangkok
Sampanth Varnakovid, Dept. of Vocational Education, Bangkok
Norani Visetbhakdi Thammasat University, Bangkok
Nanda Visissobha, Thammasat University, Bangkok
Prapajsri Vongsuradish, Dept. of Teacher Training, Bangkok
Imchai Voodhigula, Chulalongkorn University, Bangkok
Suparb Wadkhien, Chulalongkorn University, Bangkok
Udom Waratomasikkhadit, Dept. of Teacher Training, Bangkok
Pornthip Watanasuvakul, Dept. of Teacher Training, Bangkok
Cha-on Wisukhimark, Silpakorn University, Nakorn Prathom, Thailand
Sowan Wongkachorn, Dept. of Teacher Training, Bangkok
Wilaiphan Wongvanij, Thammasat University, Bangkok
Chavalit Woodhikul, Chiengmai University, Chiengmai, Thailand
Ramsamorn Yoosathaporn, Chulalongkorn University, Bangkok
Rawiwan Yuktawataya, Dept. of Vocational Education, Bangkok
Kannika Yuktiratana, Dept. of Teacher Training, Bangkok